LOCALISING FAST TRANSIENTS

A THESIS SUBMITTED TO THE UNIVERSITY OF MANCHESTER
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
IN THE FACULTY OF SCIENCE AND ENGINEERING

2019

By
Charles R. H. Walker
School of Physics and Astronomy

Contents

Word count: 52521

List of Tables

List of Figures

The University of Manchester

ABSTRACT OF THESIS submitted by Charles R. H. Walker
for the Degree of Doctor of Philosophy and entitled
"Localising Fast Transients"
September 2018.

The radio sky is not static. Time-variable radio emission is induced in the Earth's atmosphere, and is emitted by other bodies in our Solar System, the Milky Way, and in galaxies beyond our own. It may vary on timescales which span a range from years to nanoseconds. Technological advances are increasingly allowing radio astronomers to study the radio transient events which occur on the shortest of these timescales.

Of particular interest are the emission from pulsars, and an extragalactic population of fast radio bursts (FRBs), which was discovered just over a decade ago. The signals measured from both on Earth are affected by matter they traverse en-route to us. Through their dispersion measures (DMs), pulsars have enabled radio astronomers to model the Milky Way's electron distribution. Likewise, extragalactic FRBs may potentially probe matter which lies between galaxies, and that of their own hosts.

To maximise the use of such distant events, the expansion of the Universe must be accounted for. The recession of other galaxies and the phenomenon of cosmological time dilation both measurably impact observed signatures of FRBs. Both effects may be accounted for by accurate deduction of FRB source's redshifts. This is driving radio astronomers to develop radio facilities with the high sensitivities and large fields of view necessary for detection of FRBs in greater numbers, and with precise enough angular resolutions to associate them with host galaxies and thus obtain their redshifts. So far only one FRB, by virtue of its observed repeating behaviour, has been localised.

The e-MERLIN interferometer is a six-dish array of radio telescopes in the United Kingdom. Its long baselines between dishes provide it with the means to localise to sub-arcsecond accuracy any FRBs it may detect. This thesis details the Localisation of Fast Transients with e-MERLIN (LOFT-e) upgrade to e-MERLIN, developed by the author to enable high time-resolution studies with the instrument. The first successful observations of pulsars and single pulses from Rotating RAdio Transients (RRATs) are presented, techniques developed to increase its sensitivity via incoherent beamforming are detailed and the instrument's RFI environment is discussed. Finally, a framework to allow DM-redshift analysis of FRBs which may never be localised is introduced.

Declaration

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Copyright

Acknowledgements

For my family. Robert, Lesley, Toby, Kate and Roo. Guggy and Grandad. Mark Davidson. Margaret Milnes. Jim and Sue. For your encouragement, love, and generosity.

'Many a mage of great power,' he had said, *'has spent his whole life to find out the name of one single thing – one single lost or hidden name. And still the lists are not finished. Nor will they be, till the world's end.'*
-Ursula K. Le Guin

Chapter 1

Pulsars and Fast Radio Transients

1.1 The time-variable radio sky

The radio sky, encompassing electromagnetic emission from centimeter to meter wavelengths is not static (Lazio et al., 2014). Radio astronomy has existed in practice since 1934 (Jansky, 1933), and our understanding of a radio sky containing time-variable phenomena has evolved over the ensuing period. The pulsar B1919+21, discovered in 1968 (Hewish et al., 1968), is an iconic example of a periodically repeating radio source which traces its theoretical origins back to 1934 (Baade & Zwicky, 1934). It is just one example of a rich and diverse time-variable radio population which spans spacial scales from the terrestrial (e.g., atmospheric emission induced by cosmic rays; Falcke & Gorham 2003), to within our Solar System (e.g. emission from the Sun and Giant Planets; Hey 1946; Farrell et al. 1999), to the Galactic and extragalactic (e.g., flaring brown dwarf stars, pulsars and Active Galactic Nuclei (AGN) outbursts; Cordes et al. 2004a), to the cosmological (e.g., fast radio bursts (FRBs); Macquart et al. 2015.

Radio variability has been seen on timescales from sub-second to years. Some sources may appear stochastically due to one-off events: afterglows of supernovae may persist from days to decades. Some may repeat irregularly, undergoing active flaring periods, or absences of detectability. Examples include flaring stars and repeating radio transients (RRATs). Some sources, for example pulsars, are identified via periodic signals, and once identified further emission can be predicted precisely. Pulsar periods range from tens of seconds to milliseconds. Their pulse widths are intrinsically dependent on observation frequency and extrinsically dependent on temporal resolution, and are catalogued to span second to sub-millisecond timescales. Occasionally, emission from the Crab pulsar may be resolved into individual pulses nanoseconds wide (Lazio

et al., 2014; Mooley et al., 2016; Bhattacharyya et al., 2018; Hankins et al., 2016). Reviews of known sources which contribute to the dynamic radio sky may be found in, e.g., Cordes et al. (2004a); Lazio et al. (2014); Mooley et al. (2016). Such reviews may not encompass the entire radio transient population, and will be appended to as new sources are discovered.

Historically, the time-variable radio parameter space (spanning tens of orders of magnitude in both radio luminosity and event timescale) has been poorly probed due to instrumentational limitations such as observational field of view and computer processing power (Cordes et al., 2004a; Macquart et al., 2015). These limitations are being overcome as existing radio observatories are upgraded to transmit, store, and increasingly in real time, to process large volumes of data sampled at high speeds (Lazio et al., 2014). Multi-beam receivers (e.g., Staveley-Smith et al. 1996) and increasingly, phased-array-feed technology (e.g., Chippendale et al. 2015) will continue to expand the fields of view of individual dishes. More advancements and discoveries are expected with the advent of the Square Kilometre Array (SKA), which will explore hitherto unprobed regions of radio parameter space using a combination of high sensitivity, large field of view, and long observing times (Macquart et al., 2015). Its pathfinders, such as the Australian Square Kilometer Array Pathfinder (ASKAP) are already demonstrating these technologies and observing radio transients (Bannister et al., 2017).

The known population of radio transients is often separated into two classes: *slow radio transients* and *fast radio transients*. The division between these classes is potentially subjective. Mooley et al. (2016) considers phenomena behaving on timescales greater than a second to be slow transients. ASKAP, which plans dedicated surveys for both slow (VAST, Murphy et al. 2013) and fast (CRAFT, Macquart et al. 2010) transients defines fast transients as those which display variability on timescales shorter than the instrument's imaging cadence (~ 5 s). By this definition, the process of detection and analysis of fast and slow transients is performed on time-series data or images respectively, and their emission mechanisms may be separated into coherent and incoherent processes respectively (Murphy et al., 2013). Coherent emission is amplified by particles emitting in phase, and may result from stimulated emission. During incoherent emission individual particles radiate independently (Lazio et al., 2014).

Fast radio transient astronomy in particular is an increasingly exciting research field. As fast transients are associated with the most energetic events in the Universe, studying them provides insight into extreme environments associated with condensed

matter and strong gravitational fields. Additionally, the environments these signals traverse en-route to Earth impact their properties, allowing them to be used as probes of intervening media (Macquart et al., 2010). This rich scientific potential, coupled with the aforementioned technological advancements are providing the impetus to the radio transient community to resolve and analyse radio data to sub-millisecond timescales, to detect new transients at increasing rates via commensal surveys, and to consider ways to maximise the scientific yield of such events.

These objectives provide the context and the motivation for this thesis. During this PhD candidacy, the author has secured ~ 130 hours of observing time with the United Kingdom's e-MERLIN interferometer as principal investigator, leading the creation, development and use of a high time-resolution backend for the instrument. Upon completion, the *Localisation of Fast Transients with e-MERLIN* (LOFT-e) backend will enable detection, analysis, and localisation of fast transient events, using e-MERLIN's field of view, sensitivity, and sub-arcsecond localisation potential to its full advantage.

Of particular interest are the signals emitted by varieties of neutron star, and the exciting extragalactic phenomenon of fast radio bursts (FRBs), discovered just over a decade ago (Lorimer et al., 2007). The accurate localisation of cosmological FRBs is important for determination of their redshifts, which are necessary for many of their scientific purposes. However, even with recent and proposed advancements in fast transient detection, not all FRBs will be localised to sub-arcsecond accuracy, depending on their instrument of detection. The author has also considered ways to maximise the scientific yield of such sources. Both research avenues are discussed in later chapters. First, however, this chapter introduces neutron stars and fast radio bursts as the radio phenomena of interest for this thesis.

1.2 Neutron Stars and Pulsars

The existence of neutron stars was indirectly postulated in 1931 with the determination of the maximum mass of white dwarf stars (Chandrasekhar, 1931). In 1934 it was proposed that a supernova event marked the evolution of a normal star into a neutron star (Baade & Zwicky, 1934). It is now generally accepted that upon exhaustion of all fuel, tens of solar mass stars will undergo core-collapse supernova (CCSN) events (Woosley et al., 2002), leaving behind gravitationally bound cores of neutrons supported from further collapse by neutron degeneracy pressure (Özel & Freire, 2016).
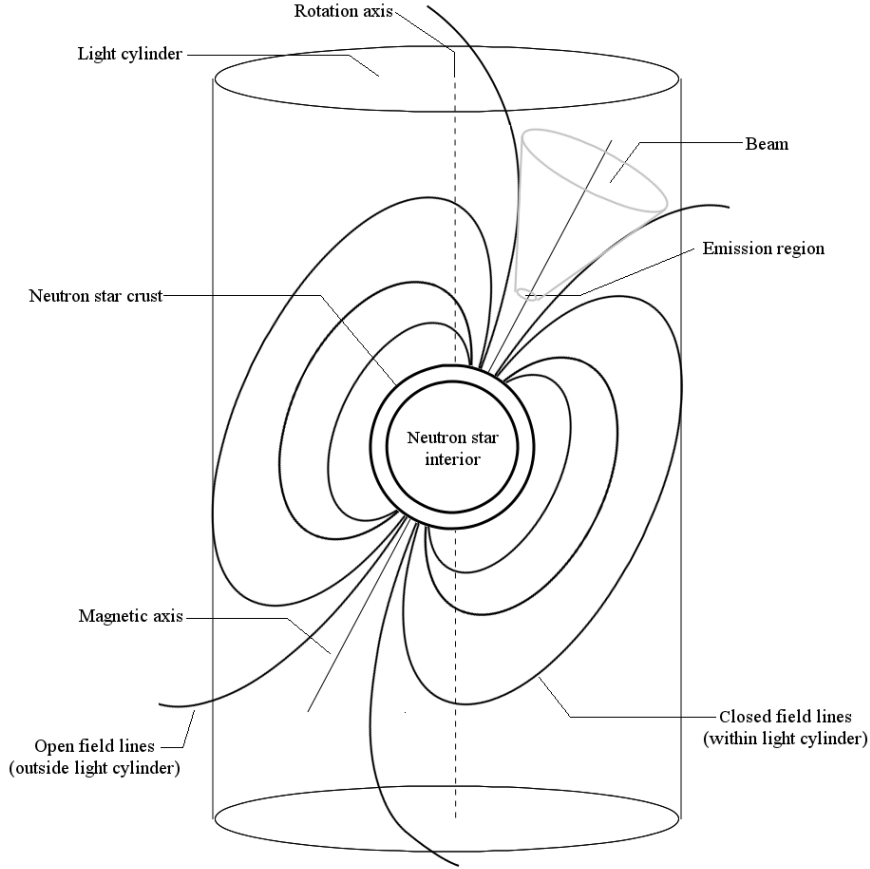
Figure 1.1: Diagram of a rotating, highly magnetic neutron star. Pulsar profiles may be explained by the lighthouse model, where charges accelerating along magnetic field lines produce a cone of radiation at each magnetic pole. If aligned appropriately with Earth, this emission will appear pulsed as it sweeps across an observer's line of sight once per rotation period. Diagram is based on that in Lorimer & Kramer (2012).

The behaviour of a gravitationally-bound degenerate gas in the absence of an energy source was discussed in Oppenheimer & Volkoff (1939), in the context of the fates of stars which had exhausted all fuel for their thermonuclear reactions. This established the idea that the properties of neutron stars may be dictated by their *equation of state*, which describes the relationship between their densities and pressures, was established. While progress has been made, the equation of state of neutron stars is not yet fully understood (Baym et al., 2018).

1.2.1   The neutron star model of pulsars

Pulsars are named for their emission behaviour, which is observed on Earth as pulses of electromagnetic radiation. It was speculated soon after their discovery that they

might be compact objects (e.g. white dwarfs or neutron stars Hewish et al. 1968). The now accepted lighthouse model, suggested independently by Pacini (1967) and Gold (1968), attributes observed pulses to a beam of radiation originating at the poles of a neutron star's magnetic axis, sweeping across the observers line of sight once every rotation period. A simplified diagram of a pulsar and its emission is shown in Fig. 1.1. Evidence to support the neutron star model includes the discovery of a pulsar within the Crab Nebula (Reifenstein et al., 1969), the remnant of a supernova observed in 1054 (Brecher et al., 1983); and its rotation rate, which implies a density ($\sim 10^{14}$ kg/cm^{-3}) exceeding the limit of degenerate electrons [1].

The composition of pulsars, their extreme rotation rates and their magnetic field strengths may be explained by way of their formation. As massive stars deplete supplies of hydrogen for nuclear fusion, it creates successively heavier elements, from helium to iron[2]. Nucleosynthesis ceases to be energetically favourable at iron[3], resulting in a core of iron where no more fusion occurs. At this point the equilibrium between outward radiation pressure and inward gravitational force which was previously held by the star fails, resulting in its collapse. If the star's core exceeds the Chandrasekhar mass $\sim 1.44\,M_\odot$, it will no longer be supported by electron degeneracy pressure, and *inverse beta decay*, whereby protons (p) and electrons (e^-) combine to form neutrons (n) and electron neutrinos (ν_e):

$$p + e^- \rightarrow n + \nu_e, \tag{1.1}$$

begins. During collapse, binding energy is lost via the electron neutrinos, infalling material impacts the core and rebounds. The core reaches the neutron degeneracy pressure, and rebounding material impacts further infalling material. A shock wave occurs, and a CCSN results (Woosley et al., 2002).

Throughout the process, conservation of angular momentum in the core:

$$I_* \omega_* = I_{NS} \omega_{NS}, \tag{1.2}$$

(for moment of inertia I and angular speed ω of a star $*$ or neutron star NS respectively) and conservation of magnetic flux Φ_B dictate that the resulting rotation rate and magnetic field strength of a post-collapse neutron star will be much greater than for its

---

[1]This may be estimated by imposing that the centrifugal acceleration for a spherical rotator be less than its gravitational acceleration at the surface to avoid breakup of the object ($\Omega^2 R < GM/R^2$).

[2]Helium, carbon, neon, oxygen, silicon, iron.

[3]Where more energy is taken to overcome the binding energy of iron than is gained by fusion.

progenitor. The neutron star model of pulsars, summarised in detail below, relies in part on these two properties (Lorimer & Kramer, 2012).

1. A simplified model of the composition of a neutron star includes an outer solid crust and an interior of superfluid free-flowing neutrons. The exact behaviour of these inner neutrons relies on the true neutron star equation of state (Oppenheimer & Volkoff, 1939; Lattimer & Prakash, 2001).

2. The star behaves as a rotating, magnetic conducting sphere. At its surface, electric forces exceed gravity and charged particles are freed to fill the pulsar magnetosphere (Goldreich & Julian, 1969).

3. Plasma, along with the pulsar's magnetic field lines, corotates with the neutron star out to radii where they would otherwise excede the speed of light (c/Ω, where $\Omega = 2\pi/P$ and P is the pulsar period). This radius, known as the *light cylinder*, separates the pulsar's open and closed magnetic field lines (see 1.1).

4. Charged particles accelerate along the open magnetic field lines, emitting tangential curvature radiation (Komesaroff, 1970).

5. A conical beam of radiation, with a width dependent on the height of the emission region along the magnetic field axis and the size of the polar cap, results. If aligned appropriately, this beam will sweep across the observer's line of sight once per rotation.

6. Depending on the angle between a pulsar's magnetic and rotation axis, and its geometry with respect to Earth, pulsars may display a single pulse (which may be formed of many sub-pulses), or a main pulse and an interpulse each rotation (see, e.g., Weltevrede & Johnston 2008 for examples interpulses).

The first pulsar, observed in 1967 by Dame Joceyln Bell Burnell at the Mullard Radio Astronomy Observatory (Hewish et al., 1968) was denoted CP 1919 (now PSR B1919+21). Its signal was observed as periodic radio emission every 1.3373 s in two receivers centered at 80.5 MHz and 81.5 MHz. There are now ~ 2600 pulsars recorded in the pulsar catalogue[4] with periods spanning five orders of magnitude, from sub-second to tens of seconds. The distribution of pulsar periods is seen in Fig. 1.2 (Manchester et al., 2005).

---

[4]`http://www.atnf.csiro.au/people/pulsar/psrcat/` (Accessed: 25/09/2018)

1.2.2 The scientific value of pulsars

Pulsars may be used as tools to estimate properties of neutron stars, such as radius and age. The latter can be estimated using a pulsar's period and period derivative. This resulting *characteristic age* relies on assumptions about the pulsar's braking mechanisim and is not always accurate (Lorimer & Kramer, 2012).

Pulsars are characterised by their so-called integrated profiles, which are formed by averaging many individual pulses. Integrated profiles are generally consistent for given observing frequencies given an appropriate amount of observing time (Lorimer & Kramer, 2012). However these profiles may exhibit a variety of behavioural changes such as *mode changes*, where features of the pulse profile vary. The integrated pulse profile provides us with a glimpse of the emission region as a pulsar beam crosses our line of sight. Models such as the fan beam (e.g., Wang et al. 2014) and hollow-cone (e.g., Komesaroff 1970) models attempt to reconcile pulsar emission models with their integrated profiles. However, as for any given source and rotation we observe a one-dimensional cut through a two-dimensional beam, and are subject to geometrical conditions of the Earth-pulsar line of sight, none has yet proven conclusive (Lorimer & Kramer, 2012).

In addition to mode changing, pulsars may exhibit other curious behaviour. Some emit giant pulses (see, e.g, Cordes et al. 2004b; Johnston et al. 2001), or additional individual pulses which do not contribute to the main pulse profile (see, e.g., Weltevrede et al. 2011). Others exhibit *nulling*, whereby radio emission occasionally becomes undetectable (Backer, 1970). In addition to (or sometimes despite their lack of) radio detection, pulsars have been seen across the electromagnetic spectrum including at gamma-ray (e.g., Abdo et al. 2013) and X-ray (e.g., Nagase 1989) energies. Studying emission properties contributes to knowledge of neutron stars and pulsar emission mechanisms.

Besides learning about neutron stars themselves, pulsars may be used as tools to probe other environments and physical regimes. They may be used to trace the density of ionised electrons in the Milky Way (e.g., Cordes & Lazio 2002; Yao et al. 2017), and its magnetic fields (e.g., Han et al. 2018). These uses are further discussed in Chapter 2. Timing pulsars has allowed tests of general relativity (see Hulse & Taylor 1975; Kramer 2016), the detection of extrasolar planets (e.g., Wolszczan & Frail 1992), and could contribute to the detection of gravitational waves (e.g., Hobbs et al. 2010). X-ray pulsars may in the future be used for spacecraft navigation (e.g., Ray et al. 2017).
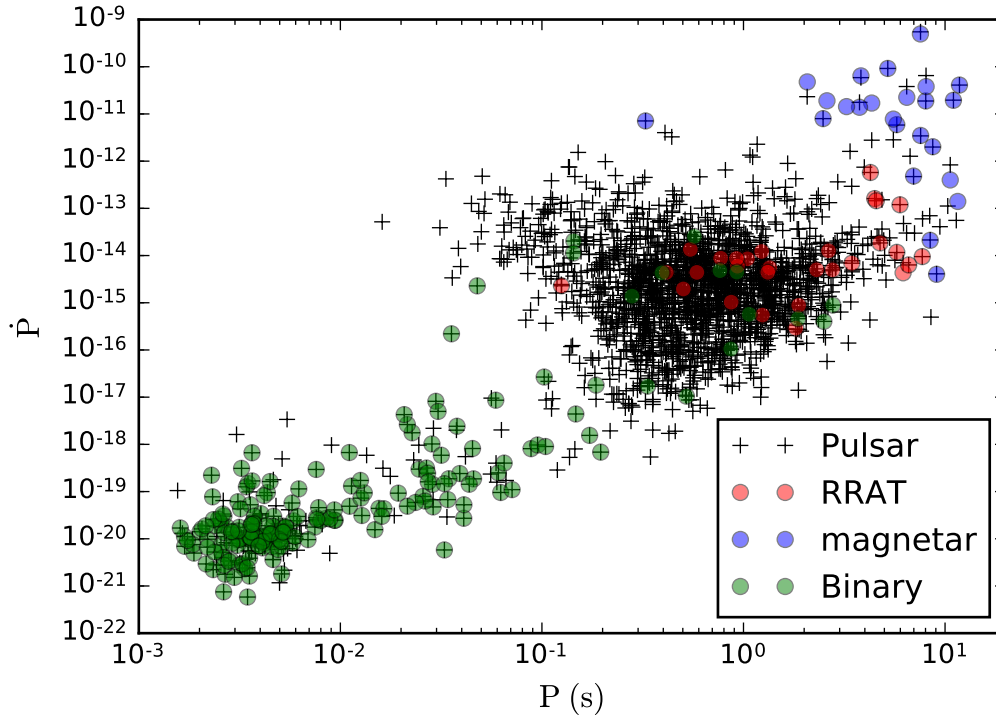
Figure 1.2: Period - Period-derivative (P-Ṗ) plot for all pulsars (black, crosses), RRATs (red, dots), and magnetars (blue, dots) catalogued to date, where both P and Ṗ were available. Green dots indicate pulsars with known binary companions. Data retrieved from the pulsar catalogue (Manchester et al., 2005) on 08/22/2018.

1.2.3   The significance of neutron star classification

Due to the emission properties of pulsars, their classification is not always straightforward. Different varieties of pulsar are of interest as they may provide insight into the evolutionary lifecycle of neutron stars. The idea that pulsars evolve with time already has precedent. An isolated neutron star's rotation rate will slow with time as it radiates energy due to a combination of particle outflow, magnetic braking and asymmetry-generated gravitational wave emission. If a neutron star forms within a binary system, it may accrete matter from its companion, spinning up and forming an *X-ray binary system* (Alpar et al., 1982). Pulsars formed in this way typically rotate with millisecond periods and are known as *millisecond pulsars* (MSPs). As seen in Figure 1.2, the vast majority (∼ 80%; Lorimer & Kramer 2012) of MSPs inhabit binary systems.

In addition to X-ray binary systems, regular rotation-powered, and intermittent pulsars, notable categories of neutron star include:

1. Radio-quiet, thermally emitting *X-ray isolated neutron stars* (XINS)[5] (Kaplan, 2008);

2. X-ray bursting *magnetars*, which are powered by extremely highly magnetic fields ($B \sim 10^{14}$ G) and are usually radio quiet (Kaspi & Beloborodov, 2017; Olausen & Kaspi, 2014)[6] ;

3. *Rotating radio transients* (RRATs). A radio source may be defined an RRAT if it is more easily detected using single pulses than by using periodicity searches[7] (Keane et al., 2011).

RRATs may be examples of weak, distant, or extremely nulling pulsars (Keane et al., 2011). The first 11 RRATs were detected in Parkes Multibeam Pulsar Survey (PMPS) data (McLaughlin et al., 2006); to date over 100 RRATs have been catalogued[8]. They emit sporadic bursts of pulses (the burst rate for an individual source can be anywhere from $0.6 - 400$ hr^{-1}). The bursts themselves have recorded flux densities $8 - 3600$ mJy and recorded widths $\sim 0.1 - 100$ ms[9]. Pulsars which exhibit additional RRAT-like bursting behaviour not related to their pulse profiles have also been discovered (Weltevrede & Johnston, 2008).

The relationship between pulsars and RRATs, magnetars and XINS is uncertain. If taken as distinct populations, the total number of RRATs in the Milky Way could exceed the total population of pulsars (McLaughlin et al., 2006; Rubio-Herrera et al., 2013). Additionally, the combined birthrates of pulsars and these classes could exceed the maximum inferred rate of CCSN[10]. See Keane & Kramer (2008) for a review, where it is suggested that a work around this *birthrate problem* is to consider the possibility that all neutron star varieties stem from the same parent population[11] .

Tentative evidence for the evolution of pulsars to RRATs and XINS exists. Where known, RRAT periods tend to be longer than those of pulsars (see the placement of RRATs on the *P-$\dot{P}$* diagram in Fig. 1.2), and XINS periods tend to be longer than those of RRATs. Temperature measurements of RRAT J1819−1458 imply it is hotter (and

---

[5]XINS are also known as X-ray dim isolated neutron stars (XDINS).

[6]`http://www.physics.mcgill.ca/~pulsar/magnetar/main.html` (Accessed: 25/09/2018)

[7]Although underlying periodicities have been detected in RRAT emission.

[8]`http://astro.phys.wvu.edu/rratalog/` (Accessed: 25/09/2018)

[9]This depends on the observing frequency.

[10]This is derived by measuring gamma-ray emission from the radioactive isotope ^{26}Al, which is created at well-known rates during CCSN.

[11]To calculate birthrate: 1) lifetime of pulsar (spin down age) $\tau = \dfrac{1}{n-1}\dfrac{P}{\dot{P}}$ where n is braking index; 2) $\tau \approx N/\beta$ where β is birthrate (Keane & Kramer, 2008).

thus younger) than the XINS population (Keane et al., 2011). It has been shown that, if placed further away, some pulsars with extremely bright single pulse emission would be identified as RRATs (Weltevrede et al., 2006).

However, relatively few RRATs have known period derivatives[12], or have been localised to allow their temperatures to be acquired. Thus the evolutionary neutron star solution to the birthrate problem has not yet been categorically confirmed (Keane & Kramer, 2008). Localising RRATs may help to expand the population of sources with these parameters constrained. This serves as an important motivation for the development of the LOFT-e backend (see Chapter 4).

1.3 Fast Radio Bursts

Pulsar surveys' long observing times, often totalling hundreds of hours on sky; fields of view often consisting of hundreds of telescope pointings; and their high temporal and spectral resolutions provide the potential for detecting phenomena other than pulsars. Such surveys have notably enabled the discovery of a fast radio transient species known as fast radio bursts (Lorimer et al., 2007).

The first FRB was discovered in archival data from a Magellanic Cloud pulsar survey (Manchester et al., 2006). While performing single-pulse searches on 480 hours of data at a centre frequency of 1.4 GHz obtained with the Parkes Radio Telescope in Australia, a single bright event less than 5 ms in duration was discovered (Lorimer et al., 2007). This event became known as the *Lorimer burst*. Its extreme properties included an estimated flux density of 30 ± 10 Jy and high dispersion measure suggestive of an extragalactic origin. Despite a subsequent 90 hours of follow-up observations around the burst's best estimated position, no repeat pulses were discovered.

A similar event, known as the *Keane burst*, was detected in Parkes Multi-beam Pulsar Survey (PMPS; D'Amico et al. 2001) data years later, again at ~ 1.4 GHz (Keane et al., 2012). Four further bursts with similar properties were detected in the High Time Resolution Universe (HTRU; Keith et al. 2010) survey dedicated to the detection of fast transients. The term FRB was coined along with these discoveries, and they were established as a population (Thornton et al., 2013). Tens of FRBs have now been catalogued[13]. While the majority of catalogued FRBs have been found at ~ 1.4 GHz, two have been found at ~ 800 MHz (Petroff et al., 2016).

---

[12]http://astro.phys.wvu.edu/rratalog/ (Accessed: 25/09/2018)
[13]Petroff et al. 2016: http://frbcat.org/ (Accessed: 25/09/2018)

FRBs have been detected by multiple telescopes worldwide. With more observatories developing the capability to detect such transients (see Section 3.3) they will be detected in increasing numbers. Recently, 309 days of observing with the Parkes telescope for the SUrvey for Pulsars and Extragalactic Radio Bursts (SUPERB) has yielded 19 FRBs, and an estimated all-sky rate of $1.7^{+1.5}_{-0.9} \times 10^3$ FRBs sky^{-1} day^{-1} with no apparent latitude dependence (Bhandari et al., 2018)[14] . Understanding of the FRB phenomenon drives the continued development of radio facilities and serves as a motivation for the development of LOFT-e.

1.3.1 The properties of FRBs

The properties of FRBs make them potentially exciting astronomical laboratories and probes of extreme physical conditions and environments beyond our own galaxy. FRBs have been distinguished from other bursting sources such as RRATs by their extragalactic origins (Keane, 2016). This is generally inferred via their *dispersion measures* (DMs). The DM is a measure of the integrated electron column density along the line of sight between an observer and a radio source and will be discussed in detail in Chapter 2.

FRBs are defined by an observed DM in excess of that which is afforded by the Milky Way along their sightlines. Therefore information regarding extragalactic environments is encoded into their DMs. Uncertainties in the models conventionally used to estimate the Milky Way's electron content and DM contribution along a sightline may have implications for categorization of a small number of FRBs and RRATs – Keane (2016), for example, identifies one RRAT with a $20 - 40\%$ chance of being a misclassified FRB due to an underestimated distance to the source – however the extragalactic origins of the majority of FRBs are beyond doubt. Figure 1.3 illustrates the stark contrast between pulsar and FRB dispersion behaviours: pulsar DMs follow the distribution of electrons in the Milky Way, while FRB DMs deviate significantly.

Other properties may shed light on physical properties of FRB sources themselves. All published FRBs are of short (durations: $\tau \sim 0.3 - 30$ ms) bright (observed peak flux densities: $S_{peak} \sim 0.2 - 120$ Jy) pulses of radio emission. As with pulsars, such short burst durations imply small emission regions for the bursts. The extreme brightness temperatures of FRBs[15] render emission processes other than coherent emission

---

[14]This rate is for FRBs within the fluence complete region. Outside this region the rate is $4.7^{+2.1}_{-1.7} \times 10^3$.

[15]See Petroff et al. 2015a for calculations of FRB brightness temperatures.
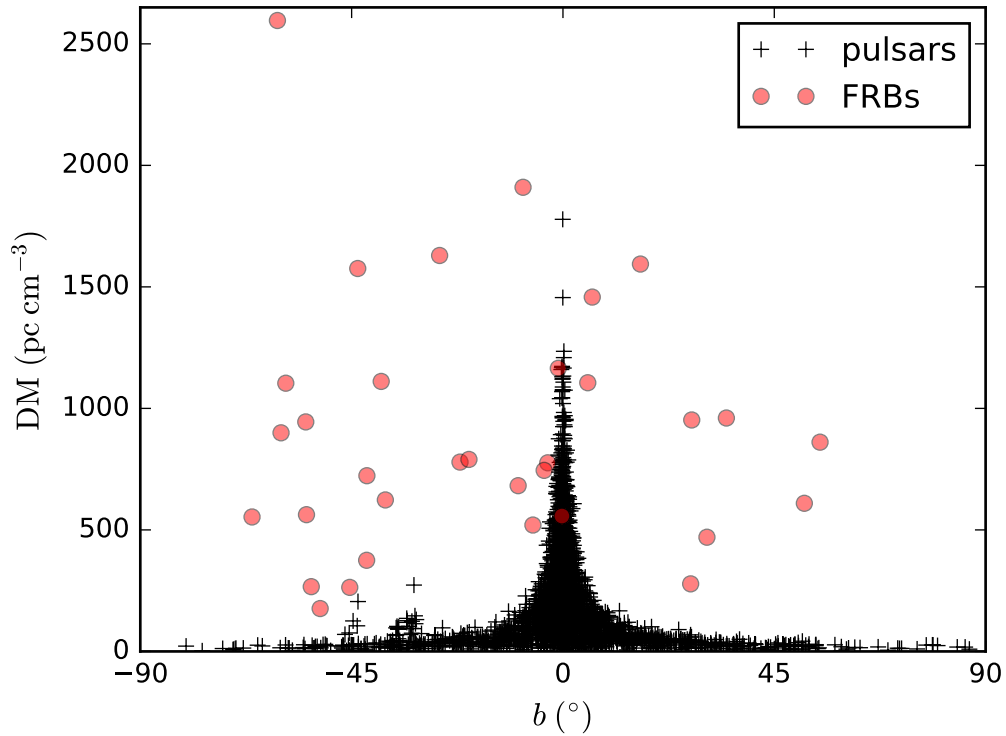
Figure 1.3: Dispersion measure (DM) vs Galactic latitude (*b*) for all catalogued pulsars (black, plusses) and FRBs (red, dots) to date. Pulsars lying within the Milky Way gain all of their dispersion from Galactic electrons, and thus will have larger DMs in denser regions of the Galaxy (e.g. closer to the Galactic center). FRBs display significant deviation from this trend: an indication that they are of extragalactic origin. Figure based on plot presented by M. Bailes at the Aspen FRB Conference 2017 (`http://aspen17.phys.wvu.edu/Bailes.pdf` (Accessed: 25/09/2018)).

unfeasible (Lyne & Smith, 1968; Katz, 2014). Macquart et al. (2015) presents an overview of the radio transient phase space, and shows that FRBs may be separated from the rest of the dynamic radio sky by properties such as their luminosities, pulse widths and observed frequencies.

An FRB's brightness or flux density S, and the related quantity fluence *F* are derived for an event detected by an instrument using the *radiometer equation*, which will be discussed in Chapter 2. The *spectral index* measures how FRB brightness depends on observation frequency, and is of significance when considering potential source emission mechanisms. To date, most FRBs have been detected by single dish telescopes. The response of telescopes is non-uniform across their beams, so the apparent brightness of an FRB will depend on where within the beam they were detected. This information is unconstrained for single dish detections (see, e.g., Obrocka et al. 2015). Therefore the observed spectral index of an FRB may not represent its intrinsic spectral index. The first observation of the only known repeating FRB yielded an apparent positive spectral index (Spitler et al., 2014). In contrast, pulsars have a spectral index of ~ -1.4 (Bates et al., 2013). Spitler et al. (2014) suggested that the detection may have been made in the side-lobe of the beam. Further burst detections from this source yielded a wide range of potential spectral indices, both positive and negative (Spitler et al., 2016). This example highlights one of the current challenges associated with determining FRB properties.

1.3.2 On the repetition and progenitors of FRBs

Only one FRB, FRB121102 (hereafter "the Repeater"), detected in Pulsar Arecibo L-band Feed Array (PALFA) data by Spitler et al. (2014) thus far displays repetitive bursting behaviour (Spitler et al., 2016). Despite multiple re-observations of fields where FRBs have been discovered including ~ 100 hrs of follow-up time for recently published SUPERB FRBs (Bhandari et al., 2018), no other repeating FRBs have been discovered. This seeming lack of repetition has been quantified by Palaniswamy et al. (2018), who conclude that the probability that no other FRBs would have been seen to repeat in all available follow-up time is between 10^{-4} and 10^{-3}, suggest therefore that the Repeater may not be representative of the entire population. Bhandari et al. (2018) further invoke differences between Repeater burst widths (which vary between $3 - 9$ ms) and those of other SUPERB FRBs, whose observed widths (< 4 ms) are only comparable due to being broadened by their high DMs (or scattered, see Chapter 2) and must therefore be intrinsically much narrower.

Theories regarding the progenitors of FRBs are diverse. They range from non-cataclysmic origins which leave the source intact (e.g. outbursting magnetars Pen & Connor 2015 or planetary pulsar companions Mottez & Zarka 2014), to cataclysmic events such as collapsing, or merging neutron stars (Falcke & Rezzolla, 2014; Totani, 2013). Repeatability can be used to constrain certain progenitor scenarios. A repeating FRB excludes any cataclysmic scenarios which fully disrupt the source. Limits on repeatability have been used to constrain the likelihood of certain progenitor scenarios, e.g. planet-pulsar systems and long period ($P < 21$ hr) repeating sources (Petroff et al., 2015c). As will be shown in Chapter 7, FRB progenitors may significantly impact the observed DM distributions of large numbers of FRBs.

1.3.3 The importance of localising FRBs

Key to both the identification of FRBs and the understanding of their nature are improved *localisation* of these sources and analysis of the environments they propogate through on their way to us. Though there is debate as to whether the Repeater is representative of all FRBs, its localisation to a host galaxy (Tendulkar et al., 2017) has allowed for intense scrutiny of the source, which may offer clues toward its nature. It lies within a star-forming region (Bassa et al., 2017b) in a low-metallicity dwarf galaxy (these have been noted to be common sources of superluminous supernovae and long-duration gamma-ray bursts; Tendulkar et al. 2017). It is also associated with a compact, persistent, variable radio source offset from the galaxy's optical center, which is postulated to be either a pulsar wind nebula, low-luminosity AGN; or a powered supernova remnant (see Marcote et al. 2017; Chatterjee et al. 2017). In addition, the extremely large and variable rotation measures (RMs) (see Section 2.4) observed for the Repeater have led to speculation that it may be a neutron star in close proximity to an accreting black hole. Michilli et al. (2018), for example, draw parallels between it and the magnetar PSR J1745$-$2900, located near Sagettarius A*, the supermassive black hole at the center of the Milky Way (Mori et al., 2013).

In addition, localisation of FRB sources allows them to be used to their maximum scientific potential. Macquart et al. (2015) reviews the uses of a cosmological population of FRBs which include: tracing 'missing' baryonic matter in the Universe which, due to their temperatures and densities cannot be accounted for by other means (see, e.g., McQuinn 2014; de Graaff et al. 2017; Tanimura et al. 2017); and offering insight into the geometry of the Universe at redshifts > 2 (Zhou et al., 2014). Many of these uses require accurate redshift measurements for FRBs (see Chapters 2 and 7 for

DM-redshift discussions.)

Due to the potential progress localising FRBs may bring, improving localisation is a key theme of this thesis. Observational techniques for searching for and localising FRBs are overviewed in Chapter 2. Techniques which could be used to improve both localisation and our understanding of progenitor and environmental factors associated with FRBs are discussed in Chapter 7.

1.4   Thesis outline

The remainder of this thesis is arranged as follows:

- **Chapter 2** describes propagation effects which affect the fast transient signals discussed in this Chapter en-route to observers on Earth.

- **Chapter 3** introduces concepts fundamental to fast transient radio astronomy, and reviews observational techniques and strategies commonly used when searching for and detecting fast radio transients. Current and upcoming facilities which will play significant roles in fast transient research in the near future are detailed to provide perspective for e-MERLIN's capabilities.

- **Chapter 4** introduces the e-MERLIN interferometer and examines its potential as a transient radio facility. LOFT-e is a new high-time-resolution upgrade to an established multi-element interferometer. Its development has been influenced by an existing framework, which is described. A particular focus is placed on the signal processing stages undertaken between initial capture at the receivers and LOFT-e's extraction point, as these may impact the form and the quality of any data received. Aquisition of the metadata necessary to make sense of any data captured is summarised, and successfully captured data are presented.

- **Chapter 5** discusses the LOFT-e data capture software and analysis pipelines developed over the course of this PhD either by the author alone, or in conjunction with others in the LOFT-e team. Techniques used during signal processing, RFI removal and searching are introduced where appropriate and discussed. Examples of their use are provided.

- **Chapter 6** presents results of the use of LOFT-e, including the first observations of pulsars and detections of RRAT single pulses by the system, and implementation of techniques to improve its sensitivity by combining multiple e-MERLIN frequency sub-bands and antennae. Planned future science utilising the backend

is overviewed, and future improvements for LOFT-e are discussed.

- **Chapter 7** discusses the implications that a cosmological origin for FRBs may
  have on their observed DMs and presents a probabilistic framework developed
  by the author to improve the scientific output of FRBs which may never be lo-
  calised to interferometric accuracy.

- Finally, **Chapter 8** reviews the major achievements of the author during this
  PhD candidacy, and the potential for future work these have generated.

Chapter 2

Propagation effects and their significance

Radio waves are directly affected by the environments through which they traverse. By the time a fast transient signal is observed on Earth its properties may have changed significantly. Thus for any given transient, information regarding, for example, ionised matter or magnetic fields along its propagation route may be directly encoded into its recorded signal. Properties such as burst or pulse duration, time of arrival as a function of frequency, intensity and polarisation state may all be affected in some way.

The nature of such effects often leaves signals more difficult to detect when unaccounted for. When quantified, however, they offer valuable insight into transient propagation environments. Study of such effects on individual sources or their populations expands their use from merely probes of the sources themselves to probes of media traversed en-route to Earth. This is a driving force behind the community's goal to detect and study as many fast transient events as possible. Galactic sources of physical emission, e.g. pulsars, have long been used to study the structure of the Milky Way. Within extragalactic transients this information is also encoded alongside further data on the intergalactic medium and the properties of host galaxies (see Chapter 7).

The next generation of fast radio transient facilities will increase the number of detected Galactic and extragalactic sources manyfold. Upgrades to existing facilities such as e-MERLIN will help contribute to this goal. With this in mind, Chapter 2 introduces prominent propagation effects which impact celestial fast radio transients. Their impact on Earth-based observations are described, and the knowledge which may be obtained from their analysis is discussed.

2.1 Dispersion and the dispersion measure

As a radio wave propagates through an ionised medium it experiences a frequency-dependent index of refraction. The result of this refractive index is *dispersion*. If a radio signal composed of a range of simultaneously emitted frequencies is dispersed, and lower frequency radiation will arrive later than higher frequency radiation.

Mathematically, the frequency-dependent time delay is written:

$$t = \mathcal{D}\frac{\text{DM}}{\nu^2}, \tag{2.1}$$

where ν is the frequency of the radiation and $\mathcal{D} = \dfrac{e^2}{2\pi m_e c}$ is the dispersion constant, derived from an electron's mass m_e, charge e, and the speed of light c. The *dispersion measure*, DM, is written:

$$\text{DM} = \int_0^d n_e \, \text{dl}, \tag{2.2}$$

where d is the distance to the source, and n_e is the electron density along the propagation sightline. Thus, the DM of a Galactic radio signal e.g. a pulsar pulse, or an extragalactic transient, e.g. an FRB, is a direct measurement of the *integrated ionised electron density along the line of sight*, i.e. the total number of electrons the signal has encountered en-route to the observer. Full derivations of both the frequency-dependent time delay and DM are provided in Lorimer & Kramer (2012).

2.1.1 Observable effects of dispersion

The effect of dispersion is demonstrated in Figure 2.1 where the characteristic quadratic sweep (Equation 2.1) of a transient pulse of radio emission is plotted as a function of frequency and time. The time delay between two frequencies in Hz, ν, as observed on Earth is given by:

$$\Delta t = \frac{1}{2.41 \times 10^{-16}} \times (\nu_1^{-2} - \nu_2^{-2}) \times \text{DM}, \tag{2.3}$$

(Manchester & Taylor, 1972).

Dispersion serves to smear out a pulse's power over time when the signal is averaged over frequency, as seen by comparing the bottom two panels of Figure 2.1. This effect renders dispersed pulses more difficult to detect during timeseries analysis. The dedispersion process (see Section 5.3 for detailed application), demonstrated in
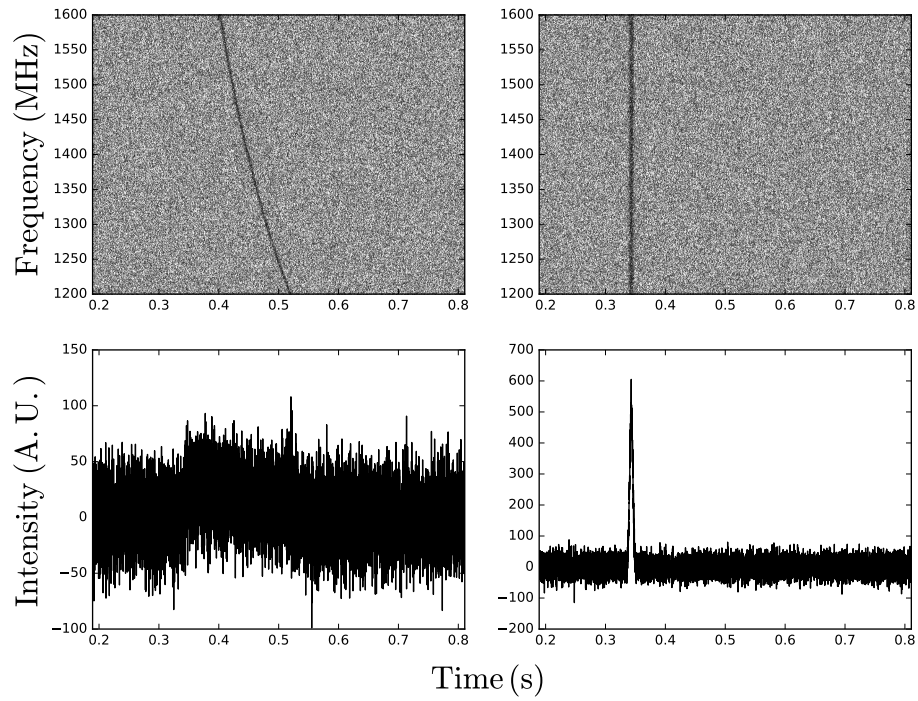
Figure 2.1: A simulated dispersed (left) and dedispersed (right) radio pulse. Top: frequency vs time dynamic spectra of the pulse. Bottom: dynamic spectra averaged over frequencies (frequency-scrunched) timeseries profiles. The greyscale corresponds to the relative flux of the simulated signal. When dedispersed correctly the pulse becomes prominent over background noise. The true DM of the pulse is $150 \, \mathrm{pc} \, \mathrm{cm}^{-3}$.

the right hand panels of Figure 2.1 is used to undo the effects of dispersion, making signals easier to detect.

2.1.2   The scientific value of dispersion

The exact form of the frequency-dependent dispersion sweeps of fast radio transients may be used to constrain properties of their emission regions and propogation environments, and separate artificial radio frequency interference (RFI) from signals with celestial origins. The kinked deviations from quadratic form identified in *perytons*, a form of local RFI previously associated with FRBs, were an early indication that they were likely of terrestrial origin (Burke-Spolaor et al., 2011). However, while the DM sweepof a burst should appear inversely quadratic in form, if the emission (or propagation) region's plasma density is large enough, higher order terms associated with the group velocity of electromagnetic waves in plasma may introduce smoother deviations from expected behaviour (see, e.g., Hassall et al. 2012). This effect has previously been used to rule out the flaring star progenitor model of Loeb et al. 2014) for FRBs by Dennison (2014). If the plasma density is high enough, the radiation will be unable to propagate at all, as an electromagnetic wave cannot travel through a region of plasma with plasma frequency, $\nu_p = \sqrt{\dfrac{e^2 n_e}{\pi m_e}}$, greater than the observing frequency. This allows upper limits to be placed on the plasma densities of transient emission regions.

Dispersion measures are also rich in scientific value. The total observed DM for a fast radio transient, DM_{obs}, may be deconstructed into Galactic and extragalactic components:

$$DM_{obs} = DM_{MW} + DM_{exc}, \tag{2.4}$$

where DM_{MW} is attributed to electrons in the Milky Way, and excess extragalactic electrons contribute to DM_{exc}. Naturally, for a Galactic source $DM_{exc} = 0$.

By combining pulsar DMs with their distances (when acquired by independent methods such as parallax), maps of the electron content of the Milky Way may be constructed. Two widely used examples are the NE2001 (Cordes & Lazio, 2002) and YMW (Yao et al., 2017) models. These electron distributions may in turn be used to estimate the distances to other pulsars which have no distance measurements. An Aitoff projection of the NE2001 model showing the maximum DM along given sightlines through the Milky Way is shown in Figure 2.2.
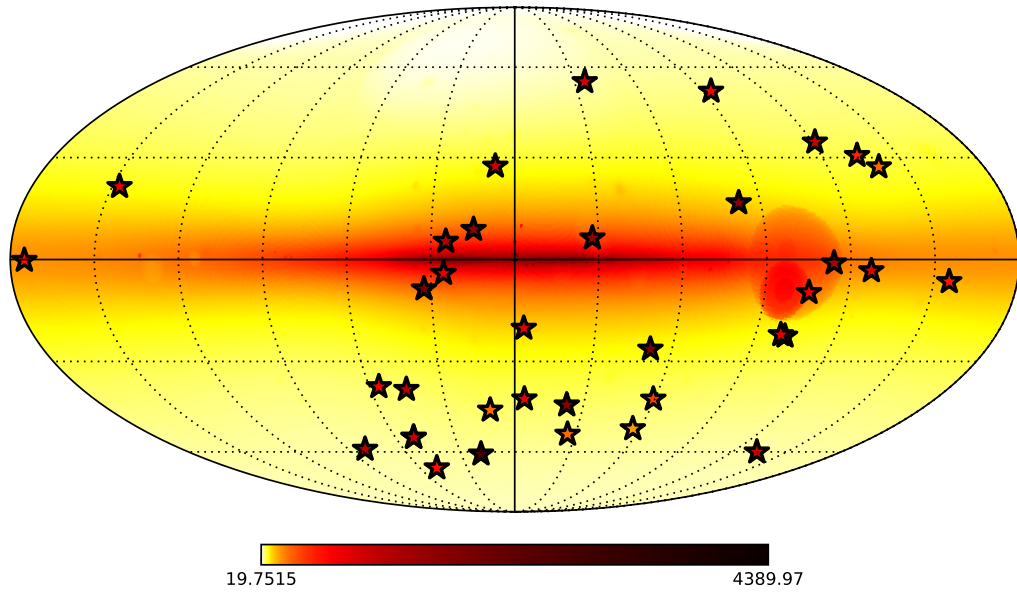
Figure 2.2: An Aitoff projection of the NE2001 model, showing the maximum dispersion measure acquired from the Milky Way along different lines of sight. Sources originating in denser regions of electrons, e.g., the Galactic plane, will have larger dispersion measures. Overplotted are all FRBs catalogued to date (stars, coloured by dispersion measure). These sources show significant excess DM, implying extragalactic origins.

It is noted that the error in DM from the NE2001 model has been estimated to be up to as much as 20% along a given sightline (Cordes & Lazio, 2002). The potential for mis-classification of RRATs as FRBs or vice-versa based on this uncertainty has been discussed in literature (see, e.g., Keane 2016). However, many FRBs have such high dispersion measures that their extragalactic origins are beyond doubt.

When the DM measured for a fast radio transient is too large to be explained by the matter in the Milky Way alone, i.e. when $DM_{exc} > 0$, the extragalactic DM contribution is usually deconstructed into:

$$DM_{exc} = DM_{IGM} + DM_{host} + DM_{local}. \tag{2.5}$$

These portions represent the DM aquired during propagation through the intergalactic medium (IGM) between galaxies DM_{IGM}, the host galaxy of the source DM_{host}, and its local burst environment DM_{local} respectively (Macquart et al., 2015). The extragalactic component introduced here is discussed in detail in Chapter 7.

For localised fast transients with known redshifts, the dependency on source redshift z_s of the IGM contribution may be used to estimate DM_{IGM}. This was demonstrated by Tendulkar et al. (2017) for the Repeater, yielding $115 < DM_{IGM} < 285 \, \text{pc cm}^{-3}$. DM_{IGM} may then be used to estimate the $DM_{host} + DM_{local}$ contribution. Such measurements, along with other observations, may provide clues for the source's progenitor. By drawing comparisons between the DMs of pulsars in the Large and Small Magellanic clouds and the Repeater's host galaxy (a star forming dwarf), Tendulkar et al. (2017) conclude that the Repeater's DM_{local} is dominated by DM_{host}. They cite this as evidence against young supernova remnant progenitors (e.g., Piro 2016) for the Repeater. Similarly, Michilli et al. (2018) use the Repeater's DM_{host} in conjunction with its observed *rotation measure* (see Section 2.4) to estimate the average magnetic field strength $< B_{\parallel}^{FRB} >$ along its line of sight. They find it to be larger ($> 0.6 - 2.4 \, \text{mG}$) than the magnetic fields found in the Milky Way ($\sim 5 \, \mu\text{G}$). From this they pose that the Repeater may be either a magnetar orbiting an accreting black hole, or a millisecond magnetar, and suggest that the magnetised environment could aid in boosting the number of observed pulses from the object.

Conversely, when the redshift of an FRB is not known due to limited localisation, the opposite approach is often taken. Estimating the $DM_{host} + DM_{local}$ contribution allows DM_{IGM}, and thus z_s to be inferred. The local and host galaxy contributions combined are usually taken to be $\sim 100 \, \text{pc cm}^{-3}$ in the host galaxy rest frame (see, e.g., Thornton et al. (2013); Lorimer et al. (2007)). Studies have shown that the mean

rest-frame host contribution could be larger ($\sim 270 \ \mathrm{pc \ cm^{-3}}$; Yang et al. 2017) and could reach up to $10^5 \ \mathrm{pc \ cm^{-3}}$ (Ioka, 2003). However, without complete knowledge of the FRB progenitor this method has limitations. By neglecting or underestimating host contributions, the derived redshift may be overestimated. In the case of the Repeater, this yields $z_s = 0.31$, compared to its true, spectroscopically obtained $z_s = 0.19$ (Petroff et al., 2016).

Galactic matter, warm intergalactic matter and X-ray corona account for $\sim 40\%$ of the Universe's baryonic matter budget as inferred by observations of the cosmic microwave background (CMB). However not all baryonic matter is easily observable (McQuinn, 2014). Recent use of the thermal Sunyaev-Zel'dovich effect (e.g., de Graaff et al. 2017; Tanimura et al. 2017) and absorption lines in quasar observations (e.g., Nicastro et al. 2018) are revealing the Universe's remaining baryonic matter. As all ionised electrons along sightlines traversed by FRBs contribute to their DMs, they may be similarly useful probes of elusive IGM baryon distributions (see, e.g., (McQuinn, 2014)), or host galaxy baryon distributions. Tendulkar et al. (2017) use estimates for $\mathrm{DM_{MW}}$ and $\mathrm{DM_{IGM}}$ along with a relation between DM and the size of ionised clouds to estimate the clumpiness of the Repeater's host galaxy ionised region.

Using $\mathrm{DM_{obs}}$ to probe the IGM becomes more robust for FRBs of larger redshifts as the contribution of $\mathrm{DM_{host}}$ and $\mathrm{DM_{local}}$ to the time delay of an FRB as measured on Earth will scale with source redshift z_s as $1/(1 + z_s)$ due to cosmological time dilation and redshift effects (Macquart et al., 2015). However, $\mathrm{DM_{host}} + \mathrm{DM_{local}}$ may still be sufficiently large as to be important depending on the host galaxy type, the location of the FRB within the galaxy, and its orientation with respect to Earth(see, e.g., (Zhou et al., 2014)). The topic of DM-redshift analysis for poorly localised extragalactic transients is explored further in Chapter 7.

2.2 Scattering

Scattering of radio transient signals is also attributed to ionised electrons along the direction of propagation. If the electrons are not distributed homogeneously, fluctuations will cause frequency-dependent multi-path propogation of the pulse due to differing refractive indices[1] .A diagram showing the effect of turbulent ionised material on the propagation of radiation from a pulsar is shown in Figure 2.3. The resulting

---

[1]The higher a signal's frequency, the less it will be deflected, and the earlier it will arrive. Consider a rainbow.
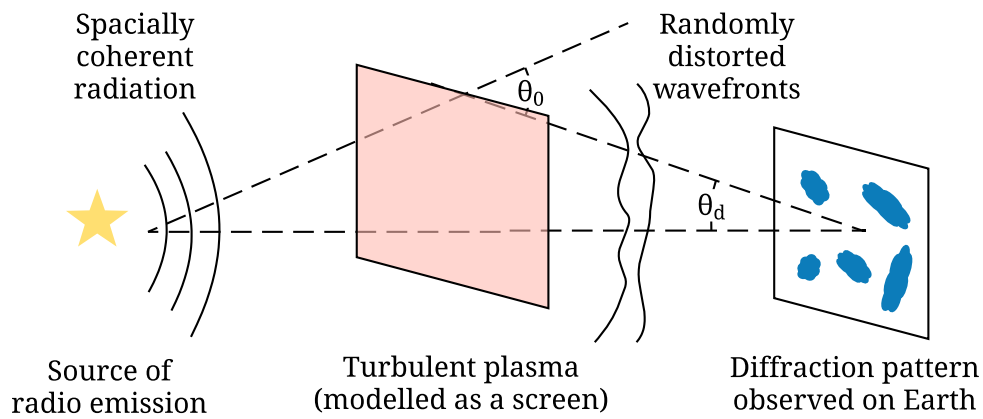
Figure 2.3: A diagram showing effects caused by inhomogeneities in the ionised electron distribution along a propagation path on radio signals . As a result of fluctuations, radiation received at angle θ_d will take longer to arrive and lead to *scattering*, frequency-dependent broadening of pulses. An interference pattern due to randomly distorted wavefronts will lead to **scintillation**, fluctuations of intensity with time. Figure modelled after from Lorimer & Kramer (2012).

frequency-dependent time delay of radiation arriving at deflected angle θ_d, known as the *scattering timescale* is given by:

$$\tau_s = \frac{\theta_d^2 d}{c}, \tag{2.6}$$

where d is the distance to the source and c is the speed of light. As $\theta_d \propto \nu^{-2}$, the effect scales strongly with frequency ν as $\tau_s \propto \nu^{-4}$ (Cordes, 2002; Lorimer & Kramer, 2012)

2.2.1 Observable effects of scattering

A simple, but well-fitting, model often used to reproduce the effect of scattering on a fast radio transient signal is the placement of a thin screen of a different refractive index to the rest of the interstellar medium (ISM) between the signal's source and the observer. The resulting asymmetric broadening of the observed pulse over time delay Δt, commonly known as a *scattering tail*, is a time-dependent intensity $I(t)'$ which can be shown to be represented as the convolution of the original pulse shape $I(t)$ with a one-sided exponential (Lorimer & Kramer, 2012):

$$I(t)' = I(t) * e^{-\Delta t/\tau_s}. \tag{2.7}$$

As with dispersion, the more a source is scattered, the more its signal is smeared out with time, leading to a decreased S/N. Therefore scattering often makes fast radio transients more difficult to detect. It has been shown (e.g. by Bhat et al. 2004) that τ_s correlates weakly with DM for a pulsar population so that higher DM pulsars are scattered more.

2.2.2 The scientific value of scattering

Some FRBs (e.g., Thornton et al. 2013; Champion et al. 2016; Bhandari et al. 2018) exhibit scattering tails which may be attributed, by analogy to the Milky Way example for pulsars, to screens in the IGM or their host galaxies. In the case of FRB 110523, detected at 800 MHz with the Green Bank Telescope (GBT), strong scattering behaviour indicative of a compact nebula or near-galactic center origin was observed (Masui et al., 2015). Thornton et al. (2013) note that if the observed scattering of FRB 110220 ocurred in the IGM, a dispersion measure due to the IGM of $DM_{IGM} > 100 \, pc \, cm^{-3}$ would be expected. The absence of scattering displayed by all perytons was further evidence that their population was distinct to that of FRBs (Petroff et al., 2015b).

2.3   Scintillation

As with scattering, scintillation is caused by inhomogeneities in the ionised medium. The effect is a fluctuation in the intensity of a radio wave as a function of both frequency and time, caused by a variable interference pattern due to randomly distorted wavefronts as shown in Figure 2.3 (Lorimer & Kramer, 2012). Mathematically, the effect is explained by phase differences, $d\phi$, which also arise from the frequency-dependent multi-path propogation of scattered pulses as discussed in Section 2.2:

$$d\phi \sim 2\pi \nu \tau_s. \tag{2.8}$$

Constructive or destructive interference of these phase-shifted signals results in the interference pattern which increases or decreases the signal's intensity with time (Lorimer & Kramer, 2012; Masui et al., 2015). Significant interference occurs when $2\pi \Delta \nu \tau_s \sim 1$, where $\Delta \nu$ is known as the *decorrolation* or *scintillation bandwidth*.

2.3.1   Observable effects of scintillation

At optical wavelengths, point sources such as stars twinkle due to turbulence in the Earth's atmosphere. Scintillation of radio waves traversing the IGM can be considered analogous to this effect. Likewise, fast radio transient signals from small sources may twinkle due to turbulence in interstellar and/or intergalactic media (Cordes, 2002; Lorimer & Kramer, 2012). The effect is variable with time depending on the relative motions of the source, scattering material and observer position (Lorimer & Kramer, 2012). This has been documented in sources such as AGN (e.g., Lovell et al. 2003) and masers (e.g., Cordes et al. 2004a) and pulsars (e.g., Narayan 1992).

Previously, Galactic scintillation has been offered as an explanation for the apparently larger numbers of observed FRBs at high Galactic latitudes via the amplification of weak off-plane FRBs (Macquart & Johnston, 2015). However, more recent studies involving larger samples of FRBs imply that this latitude discrepancy may be less statistically significant than first thought (Bhandari et al., 2018).

2.3.2   The scientific value of scintillation

In the context of extragalactic transients such as FRBs, the effect of scintillation may provide insight into their host environments and progenitors.

Alongside the aforementioned scattering (of $\tau_s \sim 1.6$ ms, see Section 2.2.2), scintillation was observed for FRB 110523 by (Masui et al., 2015), albeit with a different (microsecond) timescale. This was taken to imply the existence of two different scattering screens between Earth and the source. Similarity between the FRB scintillation behaviour and that observed for a nearby pulsar was used to attribute the FRB 110523 scintillation to the Milky Way. By imposing that the strongly scattering host galactic screen must have a smaller angular size than the screen causing the scintillation, a distance of ~ 44 kpc between the host screen and the source of the FRB was estimated.

Masui et al. (2015) used these details in conjunction with strong magnetisation (see Section 2.4) to infer that for this and similar FRBs, progenitor models involving older stellar populations (such as compact object collisions) may not support sufficiently electron-dense regions to explain the strong scattering observed. They suggested that scenarios with spinning compact objects in dense regions of ionised matter (e.g. supernova remnants or regions near galactic centers) may be preferable for such events, but noted the need for more FRBs to be detected with scintillation properties before conclusive statements can be made about the entire FRB population (Masui et al., 2015).

2.4 Faraday rotation and the rotation measure

Faraday rotation is a propagation effect undergone by a radio wave travelling through an ionised *and* magnetised medium. For a full derivation of the mathematics of Faraday rotation, see e.g, Burke & Graham-Smith (2002); Lorimer & Kramer (2012).

Electromagnetic waves consist of perpendicular electric and magnetic fields, both oscillating at right-angles to the overall propagation direction of the radiation. Radio telescopes sample the electric field of the electromagnetic wave, which can be described by a complex time-variable vector. This *electromagnetic field vector* may be written:

$$\varepsilon(A, \phi) = A \exp 2\pi i \phi(t), \tag{2.9}$$

with complex amplitude A and time-variable phase $\phi(t) = (\nu t - kz)$ for a wave of frequency ν, wavelength λ, wavenumber $k = 2\pi/\lambda$ and propagation direction z.

The position angle of an electromagnetic wave defines the direction of its linear oscillation with respect to a plane (e.g., a telescope antenna). Linear polarisation may be interpreted as the supersition of right and left-handed circularly polarised components, each with equal amplitudes A and different phases $\phi(t)$. During propagation

through an ionised and magnetised medium with magnetic field strength $B_{||}$ parallel to the path of propagation, the two components will experience different phase shifts. The Faraday rotation $\Delta\Psi_{Faraday}$ undergone by such a signal quantifies this differential shift and is written:

$$\Delta\Psi_{Faraday} = \int_0^d (k_R - k_L)dl, \tag{2.10}$$

where k_R and k_L are the right and left-handed components' wavenumbers respectively:

$$k_{R/L}(\nu) = \frac{2\pi}{c}\nu\mu = \frac{2\pi}{c}\nu\sqrt{1 - \frac{\nu_p^2}{\nu} \mp \frac{\nu_p^2\nu_B}{\nu}}. \tag{2.11}$$

Here, ν_p is the plasma frequency of the propagation medium and $\nu_B = \dfrac{eB_{||}}{2\pi m_e c}$ is the cyclotron frequency.

2.4.1   Observable effects of Faraday rotation

The rotation of the position angle of a linearly polarised wave is frequency dependent, and may be quantified for radio signals by the rotation measure (RM). The relationship between RMs and Faraday rotation is given by Lorimer & Kramer (2012):

$$RM = \frac{\Delta\Psi_{Faraday}}{2\lambda^2} = \frac{e^3}{2\pi m_e^2 c^4}\int_0^d n_e B_{||}\, dl. \tag{2.12}$$

In practice, RMs are measured in fast radio transients using frequency channels of finite width. If the frequency channels are too wide and/or if the associated RM is very large, the position angle may rotate by more than $360°$ inside a channel, the signal may be depolarised, and linear polarisation will not be measurable.

2.4.2   The scientific value of rotation measures

Pulsars are highly linearly polarised (Lyne & Smith, 1968) so their RMs and DMs may be used to measure the average magnetic field of the Milky Way between Earth and the pulsar (Lorimer & Kramer, 2012):

$$< B_{||} >= \frac{\int_0^d n_e B_\perp\, dl}{\int_0^d n_e\, dl} = 1.23\,\mu G \left(\frac{RM}{\text{rad m}^{-2}}\right)\left(\frac{DM}{\text{pc cm}^{-3}}\right). \tag{2.13}$$

FRBs are likely to be linearly polarised due to their probable coherent emission

mechanism origins (Katz, 2014; Petroff et al., 2015a). It has been suggested by Zheng et al. (2014) that of order $\sim 10^4$ FRBs with known RMs could be used to used to map the evolution of the IGM magnetic field with redshift. By drawing upon RM and redshift measurements of quasars (concluding that the IGM contribues < 30 rad m^2 to their RMs at redshifts less than ~ 0.3; Kronberg et al. 2008), they also suggest large RM and DM measurements for FRBs may indicate associated host galaxy, rather than IGM, scattering screens.

Aside from the Repeater, only four FRBs with RM measurements are recorded in the FRB catalogue to date (with values ranging from -186.1 ± 1.4 rad m^{-2} to 36 ± 52 rad m^{-2}). Some FRBs, e.g. FRB 140514, have shown circular, but no detectable linear polarisation (Petroff et al., 2015a). The authors offer three possibilities for lack of observed linear polarisation: a) that the FRB was intrinsically only circularly polarised; b) that the FRB was depolarised due to extreme Faraday rotation; and c) that the FRB was intrinsically unpolarised, the circular polarisation being induced by scintillation. These possibilities have implications for the potential progenitor and local environments of such FRBs.

Recent analysis of the Repeater by Michilli et al. (2018) indicates RMs of $\sim 10^5$ rad m^{-2}, variable by $\sim 10\%$ on half-year timescales. To measure such large RMs, Michilli et al. (2018) required data with fine frequency channels (183 kHz wide). They posed that early failure to detect linear polarisation in the Repeater (and in other FRBs) could be a combination of routine use of coarser channel widths (~ 400 kHz at 1.4 GHz) and bandwidth depolarisation arising due to the extreme $> 10^4$ rad m^{-2} RMs. This is consistent with one of the suggestions of (Petroff et al., 2015a). Michilli et al. (2018) assert that such large rotation measures may be indicative of environments close to massive accreting black holes, and draw possible comparisons between the Repeater and the magnetar PSR J1745-2900, which lies close to the MW Galactic Center and has an RM $\sim -5 \times 10^5$ rad m^{-2} (Marrone et al., 2007). A larger sample of detected FRBs with RM measurements will be required in order to establish how well the Repeater represents the entire population.

The detection of an extreme RM for the Repeater highlights the need for facilities which may process (or reprocess) fast radio transient detections flexibly. While finely channeled filterbank files may not always be ideal due to constraints (e.g. computer storage space), they are sometimes necessary to exploit observations fully. A significant advantage of the LOFT-e backend's direct access to (and the ability to record)

e-MERLIN baseband data (see Chapter 4) is the potential to create filterbanks of variable frequency channel widths as required for observations.

Chapter 3

Finding Fast Transients

As motivated in the previous chapters, fast radio transients and pulsars provide excellent tools for probing the physics of extreme objects through the baryonic content of the Universe. However, in order to achieve these scientific goals, more must be found and localised. Modern facilities (see Section 3.3) are becoming operational and have begun to detect extragalactic fast transients (Bannister et al., 2017; Boyle & CHIME/Frb Collaboration, 2018). When fully commissioned, they will supply FRBs in significantly greater numbers than ever before. Even as they do so, long established facilities may continue to contribute towards our understanding of such events by exploiting important instrumental niches. By leveraging combinations of access to high time resolution data products, large fields-of-view, high sensitivities, and the arcsecond localisation potential provided by multiple dishes, a diverse set of complimentary instruments may be established. Such is the motivation for the LOFT-e upgrade to e-MERLIN. In this chapter, relevant concepts of radio astronomy are discussed and placed into the context of fast radio transient detection. Distinctions between single and multiple-antenna instruments are made. Techniques for maximising detection rates and scientific output of events by improving sensitivity and observing field of view which are achievable by LOFT-e and e-MERLIN are discussed. Finally, prominent current and upcoming fast transient facilities and their capabilities are reviewed to provide perspective for LOFT-e's capabilities.

3.1   Fundamentals of single-dish radio astronomy

Three important factors to consider for the work discussed in this thesis are the angular resolution, sensitivity, and field-of-view (FoV) of an instrument. *Angular resolution*

defines the smallest area on the sky which an instrument may resolve, and thus its localisation potential. *Sensitivity* defines the strength of signals which the instrument may detect. *FoV* defines the total area of sky which is instantaneously visible to the instrument.

3.1.1   Angular resolution of a single dish

The relationship between the angular resolution in radians, θ, of a telescope and its operational wavelength λ is:

$$\theta \propto \frac{\lambda}{D},   \tag{3.1}$$

where D is the telescope diameter.

At radio wavelengths (which typically range from cm to m), designing telescopes with sufficiently large diameters to localise such sources becomes increasingly impractical. Consider the Five-hundred-metre Aperture Synthesis Telescope, FAST, the world's largest single dish radio telescope. FAST has an physical diameter of 500 m. The structural challenges involved with building and operating such a large instrument are immense: it is comprised of 4500 individual panels (Li & Pan, 2016), with an actual illuminated aperture diameter of 300 m. FAST's approximate angular resolution, however, is only ~ 2.4 arcminutes at a ~ 1.4 GHz observing frequency (Nan et al., 2011).

3.1.2   Sensitivity of a single dish

The sensitivity of a radio telescope is inversely proportional to its collecting area. A larger instrument has a lower threshold for minimum detectable flux density, i.e. it can detect fainter signals. The relationship between telescope size and sensitivity is quantified by its gain, G, in units of K/Jy[1]:

$$G = \frac{A_{\mathrm{e}}}{2k_{\mathrm{B}}},   \tag{3.2}$$

where k_{B} is the Boltzmann constant. Here the telescope's *effective area*, A_{e}, is used rather than its total area, as not all radiation incident on a telescope dish may reach its receiver. The effective area takes into account this inefficiency. Telescopes with larger

---

[1] 1 Jy$= 1 \times 10^{-26}$ W m^{-2}Hz^{-1}

A_e are more sensitive to faint signals[2]. The gain (G) can be used to determine how a source of flux density S_{peak} raises the noise temperature T_{peak} of a telescope:

$$S_{peak} = \frac{T_{peak}}{G},$$

(3.3)

(Lorimer & Kramer, 2012).

A figure of merit often used for quantifying the sensitivity of a radio telescope is the *System Equivalent Flux Density* (hereafter SEFD), measured in Jy:

$$\text{SEFD} = \frac{T_{sys}}{G} = \frac{2k_B T_{sys}}{A_e},$$

(3.4)

where T_{sys} is the intrinsic system noise temperature (Burke & Graham-Smith, 2002). Any instrument's effective area and noise are accounted for by its SEFD. The SEFD essentially defines the flux density of a source which, if observed at the center of a telescope's beam, would induce a noise temperature equal to T_{sys}. As such it can be used to compare the sensitivities of different instruments: more sensitive telescopes have lower SEFDs (Lorimer & Kramer, 2012)[3].

3.1.2.1 The radiometer equation

The *radiometer equation* considers the impact of the number of polarisations and bandwidth available to a reciever, and an observation's integration time, on sensitivity. The radiometer equation is used when defining the sensitivity to individual radio signals. As such it is useful for defining sensitivity to transients more easily detected via single pulse searches (e.g. FRBs or RRATs). The derivations below are found in (Lorimer & Kramer, 2012).

In the context of single-pulse detection, consider a radio burst or pulse of signal-to-noise ratio (S/N$_{peak}$):

$$\text{S/N}_{peak} = \frac{T_{peak}}{\Delta T_{int}},$$

(3.5)

where T_{peak} is the temperature of the pulse as measured by the system and ΔT_{int} are the noise fluctuations during the sampling interval of the measurement. The fundamental

---

[2]See also `https://casper.berkeley.edu/astrobaki/index.php/Radiometer_Equation_Applied_to_Telescopes` (Accessed: 25/09/2018).

[3]See also `https://www.astron.nl/radio-observatory/astronomers/lofar-imaging-capabilities-sensitivity/lofar-imaging-capabilities/lofa` (Accessed: 25/09/2018).

assumption of the radiometer equation is that for a given system the noise fluctuations behave randomly as:

$$\Delta T_{\text{int}} = \Delta T_{\text{sys}} = \frac{T_{\text{sys}}}{\sqrt{n_{\text{p}} \tau \Delta \nu}}, \qquad (3.6)$$

where T_{sys} is the system temperature, n_{p} is the number of polarisations, τ is the sampling interval of the observation, and $\Delta \nu$ is the observing bandwidth (Dicke, 1946). Substituting Equations 3.3, 3.4 and 3.6 into Equation 3.5 and introducing a *digitisation factor* which accounts for an instrument's sensitivity loss during conversion of signals from analogue voltage data to digital β[4] yields:

$$S_{\text{min}} = \beta \frac{(\text{S/N}_{\text{min}})\text{SEFD}}{\sqrt{n_{\text{p}} \tau \Delta \nu}}. \qquad (3.7)$$

Here the radiometer equation has been adapted to describe the *minimum* detectable flux density of a signal, S_{min}, with a signal-to-noise S/N_{min} by replacing S_{peak} and S/N_{peak} respectively.

When applied to periodic signals (e.g. those measured from pulsars), data are averaged over many pulses and the radiometer equation must be adapted. The assumption that receiver noise fluctuations behave randomly means that when combining data, different fluctuations add in quadrature:

$$\Delta T_{\text{int}} = \sqrt{\Delta T_{\text{sys},1}^2 + \Delta T_{\text{sys},2}^2 + \ldots} \ . \qquad (3.8)$$

During observations a pulsar's signal will only be visible for a fraction of the total observation time. As such, both on and off-pulse time on source will be observed with different fluctuations. Following Equation 3.8 they will combine as:

$$\Delta T_{\text{int}} = \sqrt{\Delta T_{\text{sys},t=t_{\text{on}}}^2 + \Delta T_{\text{sys},t=t_{\text{off}}}^2}. \qquad (3.9)$$

Expressed in terms of observation time, τ, pulse period P, and pulse width W, $t_{\text{on}} = \tau W/P$ and $t_{\text{off}} = \tau(P-W)/P$. Substituting these into Equation 3.6 and the resulting system temperatures into Equation 3.9 it follows that:

---

[4]See Section 4.2.1.1 for in-depth details on the digitisation factor β and its relevance for LOFT-e.

$$S_{\min} = \beta \frac{(\mathrm{S/N_{\min}}) \mathrm{SEFD}}{\sqrt{n_{\mathrm{p}} \tau \Delta \nu}} \sqrt{\frac{W}{P - W}}. \tag{3.10}$$

3.1.3 Field of view of a single dish

For a single pixel feed on a single-dish instrument, the FoV Ω is proportional to the angular resolution of the beamIn the simplest single-dish case, the FoV of an instrument is often quoted as a function of the telescope's half-power beam-width (HPBW)[5], which denotes the width of the beam within which half of the total power received[6] is concentrated:

$$\Omega \propto \theta^2 \propto \mathrm{HPBW}^2. \tag{3.11}$$

As the response of a telescope beam is generally close to a Gaussian with side-lobes[7], brighter signals may still be detected outside of the HPBW. The 305-m Aricebo telescope, for example, which discovered the Repeater has a HPBW of ~ 3.5 arcminutes at L-band ($\sim 1.4\,\mathrm{GHz}$[8]). Generally, the localisation of the transient within the FoV is often completely unknown. For FRBs, such events are conventionally catalogued with a position corresponding to the location of the center of the beam at the time of detection (Petroff et al., 2016).

The angular resolution dependence of Ω means that despite the increases in sensitivity and improvement in localisation potential provided by a larger dish, the resulting instantaneous FoV reduction may hinder radio transient detection numbers (Obrocka et al., 2015). Methods for increasing the FoV of a single dish instrument include the use of multi-beam recievers (e.g. the 13-beam reciever at the Parkes Observatory (Staveley-Smith et al., 1996)) or phased array feeds (PAFs, as used by, e.g., ASKAP (Heywood et al., 2016)).

Despite smaller angular resolutions, the beam sizes of even the largest single dishes are still too great to provide the sub-arcsecond localisation necessary for host-galaxy association of extragalatic transients. Alongside uncertainty regarding their true sky positions, side-lobe detections place uncertainty on properties discussed in Section

---

[5]This may also be known as beam's full width half max (FWHM), or full width half power (FWHP).

[6]Or radiated. Telescopes are essentially antennae. Antennae transmit and recieve electromagnetic radiation equally well via the repricocity principle (Neiman, 1943).

[7]Realistically. In an idealistic, 100% efficient scenario, the response would be an Airy disk.

[8]Corresponding to the commonly observed 21 cm wavelength hydrogen line, related to the hyperfine transition frequency of atomic hydrogen.

1.3.1 (Obrocka et al., 2015). However, these properties on occasion may lead to clues about their locations within the beam. The Repeater's initial discovery with an apparently positive spectral index led Spitler et al. (2014) to the suggestion of its side lobe detection.

Techniques to improve localisation of repeating sources using single dishes exist. *Gridding*, commonly used to improve localisation for pulsars, requires repeat observations of the detection area, systematically offset from the original detection pointing. A Gaussian beam of the HPBW of the instrument is fitted to the pulsar S/Ns (which will vary) recovered from each observation. From this, a more accurate position is derived. The technique is explained in detail in Morris et al. (2002). A similar technique was used to improve the localisation of the Repeater by averaging the positions of the centers of the two pointings it was re-detected in, confirming in the process that the original detection was in a side-lobe (Spitler et al., 2016).

Even if all FRBs stem from the same repeating progenitor population, the follow-up time necessary to detect more bursts may be too large, especially if repetitions are intrinsically sporadic (as has been demonstrated for Galactic sources such as RRATs). Arecibo required ~ 6 hours of follow-up time to find further Repeater pulses, but has also observed for periods of dozens of hours during which no bursts were detected. Hundreds of follow-up hours at Parkes have not yet yielded another repeater. A gridding technique which utilises multiple coherent beams or a PAF to aid localisation of a single event is discussed in Obrocka et al. (2015).

3.2   Multiple element radio astronomy

To achieve greater sensitivity and localisation potential, interferometry may be used. Interferometers consist of multiple 'elements' (e.g. antennae or dishes) used together to simulate a larger dish. The simplest case of a two-element interferometer observing a radio source is shown in Figure 3.1. The wavefront travelling at c travels different distances to each element, resulting in arrival times differing by $\Delta d/c$. This is accounted for by applying delays to the data which may be known as *geometrical delays*.

When combining signals from multiple elements, different methods may be used. These techniques offer different advantages in terms of improvement to angular resolution, sensitivity, and field-of-view. Combining multiple telescopes also offers advantages in terms of robustness to terrestrial RFI, which may hinder detection of astrophysical events (see Section 3.2.2).
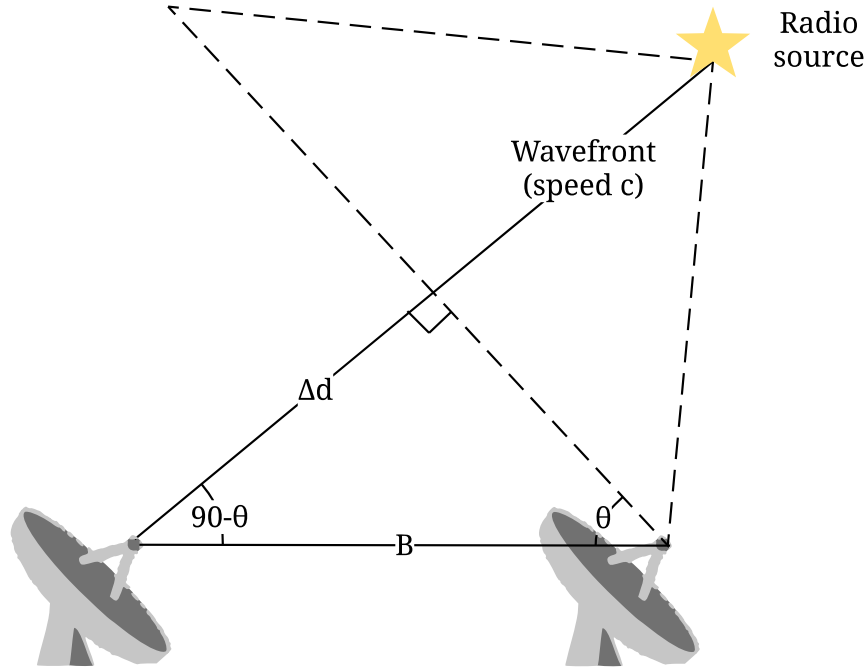
Figure 3.1: A simple two-dish interferometer observing a radio source. The difference in travel time ($\Delta d/c$) between the wavefront travelling at speed c and the two elements is accounted for using cable or digital delays, and may be known as the geometrical delay.

Different methods are subject to different limitations regarding necessary resources and available computing power. In the ideal case where resources are not limited and some (or all) methods are available to an observatory, the ideal technique may differ depending on the intended science goal, e.g. re-observation of known sources or surveying for new fast radio transients. Techniques include a *fly's-eye* approach where each element is pointed at a different patch of sky to maximise the potential FoV of the array, or *beamforming* (which may be incoherent or coherent in nature) and *imaging* techniques which focus all available elements on the same area of sky to improve sensitivity and angular resolution (see Figure 3.2). These techniques are described in detail in the following Sections. In each case the simplest configuration is considered, i.e. assuming each element of the array is identical, with diameter d_i, angular resolution θ_i, field of view Ω_i, and sensitivity $S_{\text{min},i}$. Where used, the term *baseline* refers to the distance between two elements in an array.
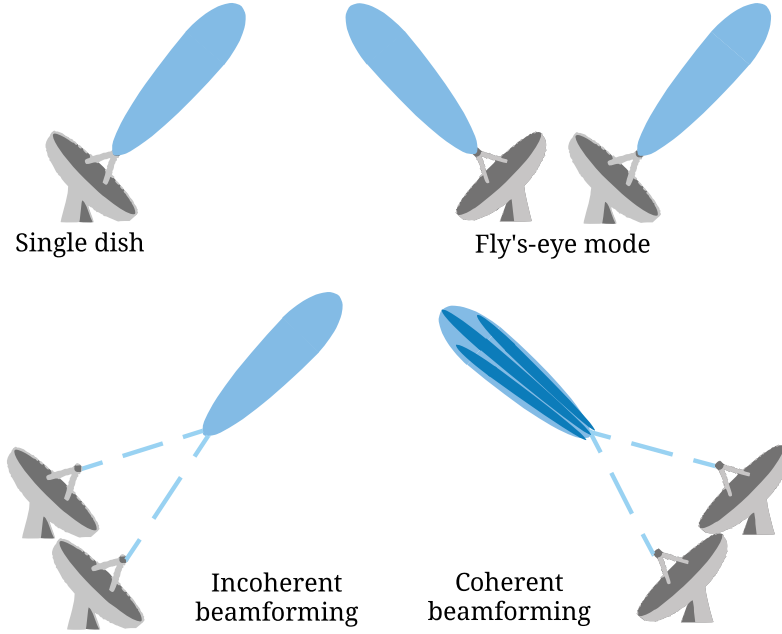
Figure 3.2: Comparisons of different observing strategies. Single dish and incoherent beam-form configurations provide equal FoV. The FoV of fly's-eye mode scales with the number of array elements. Steerable tied-array-beams with narrower fields of view may be created using coherent beamforming. Figure modelled after Colegate & Clarke 2011; Figure 2.

3.2.1   A fly's-eye approach

The fly's eye approach to element combination is conceptually the simplest. In this configuration, an array of N different elements are pointed at different patches of sky (Colegate & Clarke, 2011; Macquart, 2011).

As each element, processed separately, observes a separate area of sky containing completely different sources, the angular resolution, θ_{arr} will be constrained to that of an individual element, i.e.:

$$\theta_{arr} = \theta_i, \tag{3.12}$$

where conventionally θ_i will be quoted as the HPBW of the element. Likewise the sensitivity of the array, $S_{min,arr}$ will be equal to that of an individual element:

$$S_{min,arr} = S_{min,i}. \tag{3.13}$$

However, by pointing each element at a separate area of sky, the FoV of the array, Ω_{arr},

will increase linearly with the number of elements in the array N as:

$$\Omega_{arr} = \sum_{i=1}^{N} \Omega_i = N\Omega_i. \tag{3.14}$$

This approach maximises the observable FoV of the array without use of further equipment on each dish such as PAFs or multi-beam receivers. As such it is of potential benefit for detecting large numbers of new transients (Colegate & Clarke, 2011; Macquart, 2011).

This technique was demonstrated by Bannister et al. 2017 during a pilot survey for the Commensal Real-Time ASKAP Fast Transients survey (CRAFT) using 8 telescopes in fly's-eye mode, resulting in an FRB discovery. In this case each telescope was additionally fitted with a 36-beam PAF, increasing the total FoV by a factor of 288 over a single dish. Due to the brightness of the burst it was detected in multiple adjacent PAF beams to varying degrees of significance. The brightness of detection in these beams was used to constrain the location of the burst to a significantly smaller (by a factor of 60) area of sky than θ_i.

3.2.2 Incoherent beamforming

During incoherent beamforming, the N identical elements composing the array are pointed at the same sky location. A small difference in arrival time at each element corresponding to the different travel distance of the radio waves (see Figure 3.1) is accounted for by applying *geometric delays*. Powers[9] measured for each element of the array are then combined.

Incoherent beamforming results in an angular resolution equal to that of an individual element:

$$\theta_{arr} = \theta_i. \tag{3.15}$$

However combining signals from multiple dishes affords a sensitivity increase to observations. When observing a single pulse in N telescopes and combining them incoherently, the peak flux will sum linearly, and fluctuations will sum in quadrature. Thus for identical dishes, each with system temperature $T_{sys,i}$, Equation 3.5 becomes:

---

[9]A^2, where A is the amplitude of the signal.

$$\mathrm{S/N_{peak}} = \sum_{i=1}^{N} \frac{T_{\mathrm{peak,i}}}{\sqrt{\Delta T_{\mathrm{sys,i}}^2}}. \tag{3.16}$$

Assuming identical dishes, the effective area, $A_{\mathrm{e,i}}$ will also be the same. Following through the radiometer equation calculations as in Section 3.1.2.1 yields:

$$S_{\mathrm{min,arr}} = \beta \frac{(\mathrm{SNR_{min}})\mathrm{SEFD_i}}{\sqrt{N}\sqrt{n_{\mathrm{p}}\tau\Delta\nu}}. \tag{3.17}$$

The overall sensitivity improvement scales as $\sqrt{N}$ over an individual element. Due to the loss of the electromagnetic wave's phase information resulting from combining powers for each element, this is a lower sensitivity improvement than gained from coherent beamforming (Colegate & Clarke, 2011; Macquart, 2011).

Incoherent beamforming also results in a FoV identical to that of an individual element:

$$\Omega_{\mathrm{arr}} = \Omega_{\mathrm{i}}. \tag{3.18}$$

The tradeoff between increased sensitivity and reduced FoV afforded by incoherent beamforming vs a fly's-eye configuration should be considered by an instrument planning fast transient observations. Incoherent beamforming is beneficial for re-observing sources with known (or known to within the beam size of a single array element) locations, as sensitivity is improved while extra FoV is not required.

Macquart (2011) show that in the idealised case of a blind search for new fast transients distributed homogeneously throughout a volume of the Universe (i.e. with no latitude or redshift dependence), the expected detection rate of an incoherent beamforming survey will be $N^{1/4}$ times worse than in fly's-eye mode. They show this to be irrespective of the population's luminosity distribution (although lower-luminosity sources will be probed). However, they demonstrate that if the desired population to be probed has a location-dependent event rate[10] the favoured survey strategy for maximising detections may shift to incoherent beamforming. They show that this will depend on the size of the high event-rate region, the percentage of elements in the array needed to tile said region, and the disparity in event rates across the available survey area.

Incoherent beamforming also offers some robustness to terrestrial RFI by allowing

---

[10]See Chapter 4 for further discussion on FRB event rates.

the number of RFI-induced false positives to be reduced after a transient search. Celestial radio transients will appear in all elements of the array. However terrestrial RFI (e.g., from mobile phones or airport radar) may not be visible to dishes distributed over large areas. Thus, combinations of multiple elements may decrease the significance of local RFI in comparison to astronomical signals, and may even be used to discard certain RFI entirely depending on its strength. Additional RFI mitigation may come from coincidence techniques which discard signals seen in fewer than a threshold number of individual antennae (see, e.g., Wayth et al. 2011).

3.2.3 Coherent beamforming

As discussed in Section 1.1, the electric field vector measured by a radio telescope is a complex variable, consisting of an amplitude A and a phase $\phi(t)$. During incoherent beamforming, only the power (A^2) is considered, and phase information is lost. During the coherent beamforming process, phase shifts are applied individually to each element of an array, steering and aligning them completely on a patch of sky. Voltages are then summed. In addition to geometric delays, *phase delays* accounting for ionospheric effects along a particular direction on sky are applied (Colegate & Clarke, 2011). The required phase delays will vary as a function of time over timescales similar to those of changing atmospheric fluctuations, which are dependent on baseline and observing frequency and will vary across the sky. Thus, phase corrections are calculated for a point source close to the intended target, known as a *phase calibrator*. These may then be applied to observations of the target.

Coherent beamforming results in the formation of a *tied-array-beam* (TAB) with a narrow FoV. For an array of longest baseline B, composed of identical dishes of diameter d_i and angular resolution θ_i, the angular resolution of the TAB scales as:

$$\theta_{\text{TAB}} = \theta_i \times \left(\frac{d_i}{B}\right) \propto \frac{\lambda}{B}, \tag{3.19}$$

and the localisation potential of a single TAB is therefore proportional to the HPBW of the TAB (Obrocka et al., 2015). Multiple TABs may be created and this may be exploited. Obrocka et al. (2015) discuss techniques utilising adjacent TABs to achieve accuracies greater than a single HPBW.

The sensitivity increase afforded by a coherent beam is a factor N improvement from the individual element case and a factor of $\sqrt{N}$ improvement from the incoherent beam:

$$S_{\text{min,arr}} = \beta \frac{(\text{SNR}_{\text{min}})\text{SEFD}_{\text{i}}}{N\sqrt{n_{\text{p}}\tau\Delta\nu}}. \tag{3.20}$$

(Colegate & Clarke, 2011; Obrocka et al., 2015). The FoV of a single TAB is proportional to its angular resolution, which will generally be quoted in terms of its HPBW:

$$\Omega_{\text{TAB}} \propto \theta^2_{\text{TAB}} \propto \text{HPBW}^2_{\text{TAB}}. \tag{3.21}$$

Given unlimited resources, multiple TABs may be created to tile the array's incoherent beam FoV, maximising both the sensitivity and FoV of the array. In practice it becomes computationally expensive to search many TABs for transients, thus the formed will be resource limited. For an array creating N_{TAB} TABs, the FoV will increase linearly with N_{TAB}:

$$\Omega_{\text{arr}} = N_{\text{TAB}}\Omega_{\text{TAB}}, \tag{3.22}$$

(Colegate & Clarke, 2011).

When re-observing known sources with locations known to within Ω_{TAB}, coherent beamforming is preferable to incoherent beamforming or a fly's-eye configuration due to its increased sensitivity.

As discussed in Section 2.2.1, scattering due to the inhomogeneous ISM can bias against short-period pulsar detections in Galactic pulsar surveys, which may lead to direction-dependent event rates (Lyne et al., 1998). This may also be true for other fast radio transients in Galactic surveys. Macquart (2011) show that in specific cases, a Galactic survey using an N-element array in a coherent beamforming configuration may yield fewer events than if using an N element fly's-eye approach, and that coherent beamforming becomes less effective with increasing N. This is heavily dependent on the Milky Way's geometry, scattering effects, and thus the chosen survey field. However, they show that at sufficiently large N, when $S_{\text{min,arr}}$ becomes able to probe outside the Milky Way and the survey becomes sensitivity-limited, this will no longer be the case. Extragalactic, sensitivity-limited surveys (if the fast radio transients in question are homogenously distributed) expect detection rates of $R \propto \Omega_{\text{arr}}S_{\text{min,arr}}^{-3/2}$. In the most extreme case where $S_{\text{min,arr}}$ is low enough to probe to the earliest progenitors of the sources, the survey becomes volume limited. Macquart (2011) note that for sensitivity or volume-limited surveys, coherent beamforming will be the optimum strategy if enough TABs can be created and processed to account for their smaller FoV. Coherent beamforming is also more resistant to RFI than incoherent beamforming, as RFI

should decorrolate.

3.2.4 Imaging

During interferometric observations of slowly varying radio sources, imaging techniques are commonly used. During imaging, sources are identified and studied in the image plane rather than in timeseries data.

Interferometric arrays typically perform imaging using a dedicated correlator. During imaging, phase calibrated data are correlated for each element in the array, averaged over an integration period, and *complex visibilities*, which compare the phases between elements, are formed. These visibilities are gridded to form a UV plane, which is Fourier transformed to make a pixellised image of the sky (Colegate & Clarke, 2011)[11].

The classification of a 'fast' vs 'slow' transient is potentially subjective. Sources are often defined based the timescale of their variability in comparison to the imaging timescale of the instruments of their detection. Traditionally correlators will integrate images over timescales > 1 s (Law et al., 2011). ASKAP, for example, has an image cadence of ~ 5 s and has commissioned its fast radio transient (CRAFT; Macquart et al. 2010) and slow radio transient (VAST; Murphy et al. 2013) surveys accordingly.

Typically, fast radio transient analysis is performed on individual beams as long correlator integration times may impair fast transient discovery smearing out variability, or inhibit measurement of properties such as dispersion (Colegate & Clarke, 2011; Law et al., 2011; Obrocka et al., 2015). However 'fast' correlator technology capable of supplying visibilities on timescales < 1 s has been developed, demonstrated detection and localisation of pulsar giant pulses in the image plane (Law et al., 2011), and used to search for FRBs with the Very Large Array (VLA) (Law et al., 2015). Potential advantages and disadvantages of fast imaging vs coherent beamforming have been found to be dependent on the specifics[12] of the instrument in question (Colegate & Clarke, 2011)[13].

As with coherent beamforming, the angular resolution acheivable by an array of N elements using imaging techniques is proportional to the largest baseline B between its elements:

---

[11]`https://www.skatelescope.org/uploaded/55809_97_memo_Cordes.pdf` (Accessed: 25/09/2018)

[12]E.g., system temperature.

[13]See also Colegate & Clarke 2011 citations, e.g.: `https://www.skatelescope.org/uploaded/55809_97_memo_Cordes.pdf` (Accessed: 25/09/2018).

$$\theta_{\mathrm{arr}} \propto \frac{\lambda}{\mathrm{B}}. \tag{3.23}$$

This angular resolution forms the dimensions of pixels in the gridded image and quantifies the localisation potential of an array.

Likewise, as with coherent beamforming, by conserving both amplitude and phase information of signals, imaging theoretically achieves the maximum possible sensitivity for the array:

$$S_{\mathrm{min,arr}} = \beta \frac{(\mathrm{S/N_{min}})\mathrm{SEFD_i}}{\mathrm{N}\sqrt{n_{\mathrm{p}}\tau\Delta\nu}}, \tag{3.24}$$

and during imaging the entire field of view of an individual array element is retained:

$$\Omega_{\mathrm{arr}} = \Omega_{\mathrm{i}}. \tag{3.25}$$

The first interferometric observations of FRBs were made by Caleb et al. (2017) with the Molonglo Observatory Synthesis Telescope using beamforming techniques, and localised to 15 arcsec × 8.4 degrees accuracy. All other FRBs have also been initially detected using beamforming. However, advances in real-time detection at interferometric facilities are making retroactive imaging of detected FRBs possible. By dumping visibilities integrated over sub-millisecond timescales using the VLA's fast correlator[14], sub-arcsecond interferometric localisation of the Repeater, sufficient for host galaxy identification, has been achieved (Chatterjee et al., 2017). Additionally, the VLBA's V-FASTR experiment (Wayth et al., 2011) searches for transients commensally in real-time and stores raw voltage data around the time of significant events. It aims to use the software correlator DiFX (Deller et al., 2007, 2011), for retroactive correlation, imaging, and subarcsecond localisation of extragalactic fast radio transients.

3.3   Current and upcoming radio transient facilities

Fast transient radio astronomy is a rapidly developing field. This is in part due to the wealth of scientific knowledge still to be obtained by studying such events, as

---

[14]Referred to in Law et al. 2015, and in Chatterjee et al. 2017 by the name Realfast.

discussed in Chapters 1 and 2. It is also due to improvements in computing and processing power using graphical processing units (GPUs). Rapid development of efficient data processing, dedispersion, and fast transient search techniques (e.g., Barsdell et al. 2010, 2012a,b; Bassa et al. 2017a) now allow events previously detected retroactively to be found and studied in real-time. This has led to the ability to perform high temporal and frequency resolution analysis via recording of raw voltages, uncovering new features in transients such as FRBs (e.g., Farah et al. 2018). Opportunities to broaden the scope of fast transient research via collaborative multi-wavelength and multi-messenger follow-up have also arisen thanks to these developments (Keane et al., 2014; Andreoni & Cooke, 2018).

This motivates the development of the LOFT-e backend for e-MERLIN as a complimentary instrument to currently established and upcoming fast transient research facilities. To provide perspective for LOFT-e's capabilities, Table 3.1 summarises important fast radio transient specific aspects of some of the major facilities in the research field.

The factors listed for each facility include:

- **Operating center frequency** ν and **operating bandwidth** $\Delta\nu$. Wide bandwidths over a range of frequencies are important for study of spectral behaviour, spectral index and scintillation. Bandwidth dictates the minimum sampling time of the facility (see Chapter 4). Fast sampling times allow study of high resolution temporal behaviour via coherent dedispersion (see Chapter 5).

- **Minimum angular resolution** θ. Where possible, the HPBW is used. Small angular resolutions dictate localisation potential, and enable maximum scientific output for extragalactic fast radio transients.

- **Minimum detectable flux density** S_{min}. More sensitive instruments will detect more events. However for facilities which are capable of different modes, care must be taken when trading sensitivity vs FoV. Where provided, the pulse width τ and S/N threshold used to calculate S_{min} have been included.

- **Maximum field of view** Ω. Large FoVs are useful for detection of many events and maximising survey speeds.

- **Elements** of telescope. A telescope may be single element or multi-element. For multi-element arrays, the observation mode (as discussed in previous sections) dictates the sensitivity, FoV and angular resolution available to the instrument. Some multi-element arrays may additionally include PAFs which increase their FoV.

- **Voltage dump ability**. The ability to dump raw voltages enables potential offline localisation for interferometric arrays and enables maximum temporal resolution to be reached.

- **Stokes parameters** available to the facility. By combining polarisations to calculate Stokes parameters, polarisation properties of events may be studied.

Telescope (survey)	ν MHz	$\Delta\nu$ MHz	θ	S_{min} Jy	Ω	Elements	Voltage dump? Yes/No	Stokes	Ref.
Parkes (SUPERB)	1382	~ 340	14.0'	0.025 ($\tau \sim$ 250 ms)	0.55 deg^2	Single	No	Full	1,2
Aricebo (PALFA/PUPPI)	1375	300	3.5'	0.065 ($\tau = 1$ ms S/N=10)	0.022 deg^2	Single	Yes	Full	3,4
Molonglo (UTMOST)	843	31.25	15"	1.6 ($\tau = 1$ ms S/N=10)	~ 16 arcsec^2	Multi	Yes	Full	5, 6, 7
VLBA (V-FASTR)	1550	64	<1"	0.27 ($\tau = 1$ ms S/N=1)	0.27 deg^2	Multi	Yes	Full	8, 9
CHIME (CHIME/FRB)	600	400	~ 2'	-	250 deg^2	Multi	Yes	I	10, 11, 12
ASKAP (CRAFT)	1320	336	8'	~ 20 ($\tau = 1$ ms)	160 deg^2	Multi	Yes	Full	13, 14, 15
MEERKAT (MEERTRAP)	1300	750	~ 10'	~ 0.001 ($\tau = 1$ ms S/N=10)	1.27 deg^2	Multi	Yes	Full	16
WSRT (ALERT)	1425	300	~ 25'	0.46 ($\tau = 5$ ms S/N=10)	8.7 deg^2	Single (Multi)	-	I	17

Table 3.1: List of fast radio transient facilities and their capabilities. Note: SUPERB discards a portion of overall bandwidth due to RFI. Note: CRAFT values demonstrated during FRB detection. Note: ALERT is a collaboration between WSRT and LOFAR. References: 1 (Keane et al., 2018) 2 (Bhandari et al., 2018) 3 (Spitler et al., 2014) 4 (Michilli et al., 2018) 5 (Caleb et al., 2017) 6 (Farah et al., 2018) 7 C. Flynn (via priv. com.) 8 (Wayth et al., 2011), 9 (Burke-Spolaor et al., 2016) 10 (Ng et al., 2017), 11(The CHIME/FRB Collaboration et al., 2018), 12 C. Ng (via priv. com.)13 (Bannister et al., 2017) 14 (Macquart et al., 2010) 15 (Clarke et al., 2014) 16 `http://caastro.org/wp-content/uploads/2018/06/Caleb-FRB2018.pdf` 17 (Maan & van Leeuwen, 2017) (Accessed: 25/09/2018)

Chapter 4

Localisation of Fast Transients with e-MERLIN

Chapters 1, 2, and 3, introduced fast radio transient phenomena and the scientific value of their study, and discussed concepts of radio astronomy relevant to the detection, study, and localisation of pulsars and radio transients. These introductory chapters provide context for the main body of this work, which introduces a new high time-resolution backend for the UK's six-dish[1] e-MERLIN interferometer.

It is clear that over the decade since the discovery of the first fast radio burst, fast transient astronomy has undergone rapid development. The next generation of fast radio transient search facilities are currently undergoing commissioning. Some are already detecting transients and offering new insights into their natures and behaviour. Many current contributors to the field are using established facilities that have undergone upgrades to increase their observational bandwidths, improve temporal and spectral resolutions, and maximise localisation potential. Some facilities are still being upgraded. Emphasis is being placed on the ability to analyse data commensally to regular observations in real time, and on facilitating collaboration between instruments around the world to maximise scientific output.

The potential contribution that extragalactic transients may add to our knowledge of the Universe has also inspired the development of LOFT-e, a high time-resolution backend for e-MERLIN. Exploiting the interferometer's multiple radio telescopes, long baselines, and its ability to store high time-resolution data for multiple polarisations at multiple observational frequencies expands the research phase space available to the observatory. Its ability to achieve the sub-arcsecond angular resolution necessary

---

[1]Nominally, although the Lovell Telescope may also be added to the array.

for localising extragalactic transients, allows it to potentially maximise the scientific yield of any events it detects.

In this chapter the e-MERLIN interferometer is introduced and its technical specifications are reviewed. The instrument's potential as a fast radio transient observatory is discussed. As e-MERLIN is an established facility, LOFT-e has been designed to exploit existing hardware and pre-existing frameworks for data and metadata acquisition. The associated stages of processing undergone by data en-route to LOFT-e inevitably impact its results. In some cases these stages are still being commisioned and LOFT-e's feedback to e-MERLIN staff has contributed to the facility's overall development. Thus the signal chain between e-MERLIN telescopes and LOFT-e is described here and placed into context of the project. An overview of LOFT-e's hardware which captures all necessary data and metadata is provided. Finally, an example of e-MERLIN data successfully captured by LOFT-e is presented.
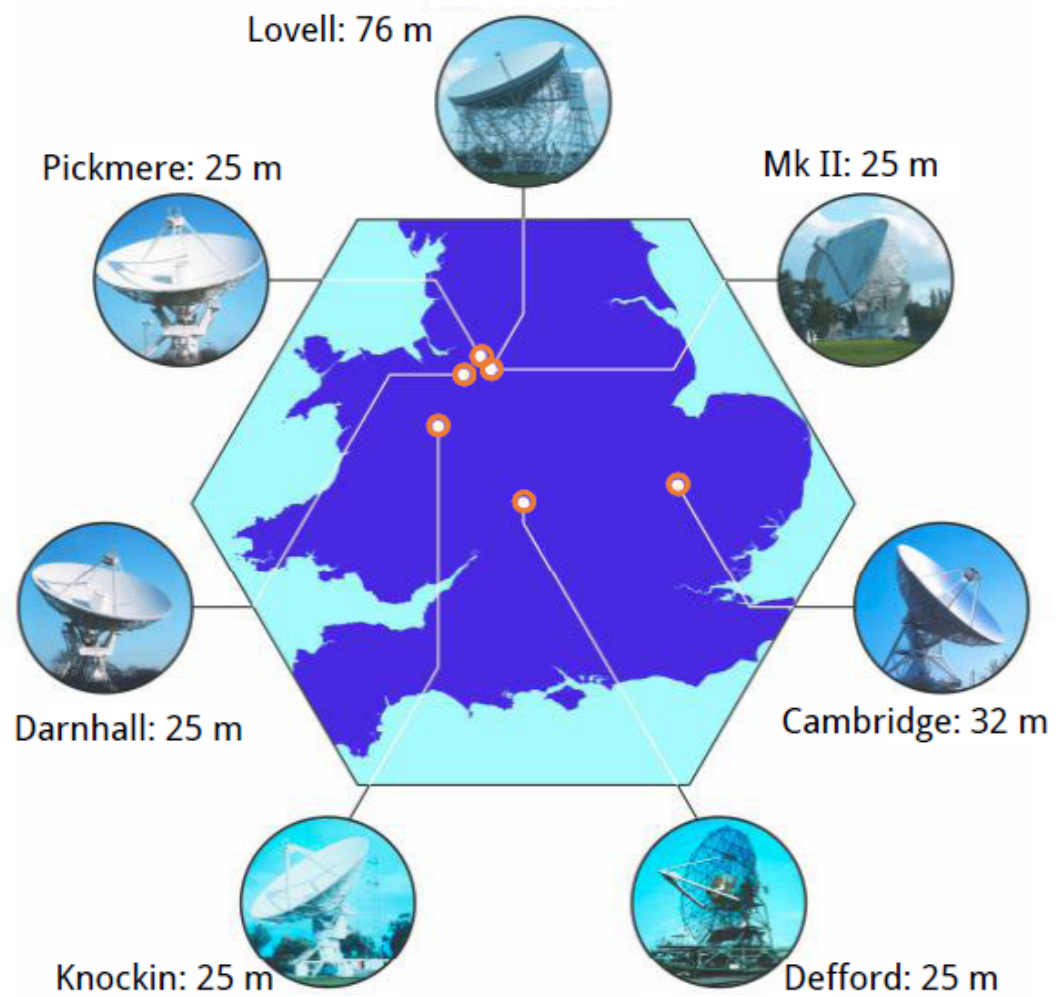
Figure 4.1: Arrangement of e-MERLIN's antennae. Image source: `http://www.merlin.ac.uk/` Image licensed under the Creative Commons Attribution 3.0 unported license (`https://commons.wikimedia.org/wiki/File:MERLIN_map.jpg`) Accessed: 25/09/2018.

Telescope	Diameter (m)	SEFD (Jy)
Mark II	25*	500
Pickmere	25	500-650
Knockin	25	500-650
Darnhall	25	500-650
Defford	25	500
Cambridge	32	250-300
Lovell	76	50

Table 4.1: Specifications of e-MERLIN telescopes as of 23/08/2018. SEFDs were provided by R. Beswick via private communication. SEFD ranges cover, instrumental, recieving band, and elevation variations. *The quoted diameter of the Mark II is an *equivalent diameter* (see text)[2].

4.1   The e-MERLIN interferometer

The e-MERLIN interferometer is a UK-based radio telescope and an upgrade to the *Multi-Element Radio LINKED interferometer* (MERLIN)[3]. The array nominally comprises six dishes distributed across the UK. The longest baseline between two dishes is 217 km (Garrington et al., 2004; Spencer, 2009; Garrington & Beswick, 2016).

The smallest four e-MERLIN telescopes, situated near the villages of Darnhall, Defford, Knockin and Pickmere are circular dishes, each of diameter 25 m. The Mark II telescope at Jodrell Bank Observatory is an elliptical dish with semi-major and semi-minor axes of 38.1 m and 25.4 m respectively (Lovell, 1964), but is generally assumed to be the equivalent of a 25 m dish[4] The sixth dish, near Cambridge, is 32 m in diameter. Optionally the 76-m-diameter Lovell Telescope, may be added to the array to increase its sensitivity at the expense of its FoV (Garrington & Beswick, 2016). The individual dishes and their distribution across the UK are depicted in Figure 4.1. System equivalent flux densities for each telescope at time of writing are provided in Table 4.1.

---

[2]See, e.g., `http://www.e-merlin.ac.uk/tech/` (Accessed: 25/09/2018).

[3]The upgrade took place between 2009 and 2012. Alongside improvements to the telescope's recievers and correlator, a network of fibre-optic cables was utilised to increase the instrument's operational bandwidth from 14 MHz per polarisation to current values (Spencer, 2009; Garrington & Beswick, 2016).

[4]See, e.g., `http://www.e-merlin.ac.uk/tech/` (Accessed: 25/09/2018). The approximation of the Mark II as a 25 m dish will be made throughout the remainder of this thesis.

The e-MERLIN interferometer may operate at three different observational frequency bands, referred to as L-band (1.3 GHz - 1.8 GHz), C-band (4 GHz - 8 GHz) and K-band (22 GHz - 24 GHz). An annual ~ 4250 hrs of observing time are distributed between these bands, in a planned 40%:40%:20% L:C:K-band ratio. For a single observation, e-MERLIN is at present hardware constrained to providing 512 MHz bandwidth at all times.

Half of the facility's observing time is allocated to *e-MERLIN Legacy Program* projects (Spencer, 2009). These projects cover a diverse range of astronomical interests from massive star formation to galaxy evolution (see Garrington & Beswick (2016) for a comprehensive summary of e-MERLIN's legacy projects). The remaining observing time is allocated to the worldwide astronomical community by way of research proposal submissions[5]. Calls for proposals are issued semi-annually. These proposals are peer-reviewed, and portions of the available observing time for the facility are allocated accordingly.

Currently, applicants of successful projects are issued correlated visibility data products[6]. The fastest visibilites available are integrated over 1 s timescales[7] Theoretically, visibilities averaged over 250 ms timescales are achievable by the correlator, however e-MERLIN backend hardware and software constraints prevent the use of this ability.

The e-MERLIN correlator is capable of outputting 128 MHz sampled, dual-polarisation voltage data in tandem with its visibilities. This is achieved by leveraging an interface designed for e-MERLIN to e-VLBI connectivity (Spencer, 2009). Currently, no higher time-resolution data products are offered as part of routine operations, and no high time-resolution operating mode exists. Since 2014, the author has successfully secured observing time via proposal calls for the development of LOFT-e to exploit this capability. Upon completion, LOFT-e will provide access to high time-resolution data products to e-MERLIN users for fast radio transient research.

4.1.1 A potential fast radio transient facility

Multiple dishes provide e-MERLIN with the potential sensitivity necessary to detect fast radio transients. Although most FRB detections are expected to be made with the

---

[5]Note: ~ 72 hours per year are reserved for development of e-MERLIN - VLBI integration.

[6]See, e.g., e-MERLIN's most recent observing call for full e-MERLIN proposal details at time of writing: `http://www.e-merlin.ac.uk/observe/call_cycle7.html` (Accessed: 25/09/2018).

[7]This timescale is consistent with most modern traditional correlators (see Section 3.2.4 for a summary).

Telescope Configuration	$\theta_{\text{e-MERLIN}}$	$S_{\text{min,e-MERLIN}}$	$\Omega_{\text{e-MERLIN}}$
Single dish	~ 35 arcmin	~ 11.9 Jy	~ 0.27 deg^2
Fly's-eye mode	~ 35 arcmin	~ 11.9 Jy	~ 1.63 deg^2
Incoherent beamforming	~ 35 arcmin	~ 4.8 Jy	~ 0.27 deg^2
Coherent beamforming	~ 0.25 arcsec	~ 2.0 Jy	~ 0.28 arcsec^2

Table 4.2: Summary of e-MERLIN's potential capabilities in various observing configurations. For simplicity, all calculations approximate the array to six identical 25 m dishes of SEFD 600 Jy. Sensitivities are calculated for a single pulse fast radio transient event of width $\tau = 1$ ms and SNR $= 10$ over a bandwidth $\Delta\nu = 128$ MHz and with a digitisation factor $\beta = 1$. All calculations are for L-band (1.4 GHz). Coherent beamforming FoV is for a single beam.
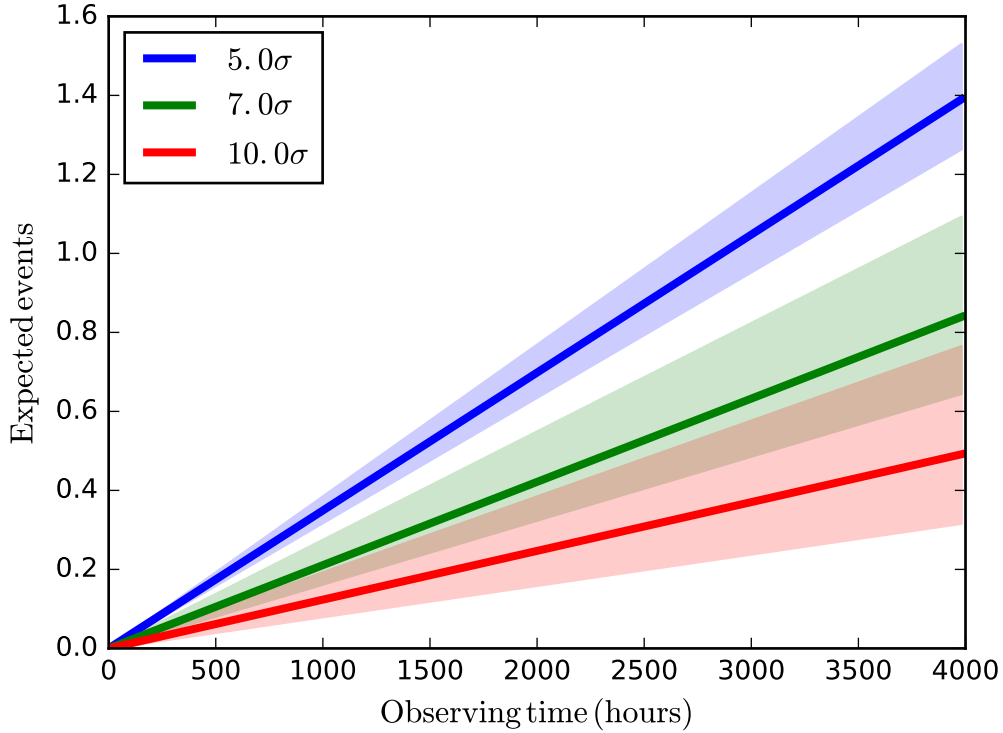
Figure 4.2: Expected number of FRB events detected by a six telescope e-MERLIN incoherent beam (for 128 MHz observing bandwidth at L-band) as a function of observing time. Coloured lines correspond to different burst signal to noise thresholds assuming a power-law flux intensity distribution with index $\alpha = 3/2$. The shaded regions indicate boundaries for $\alpha = 1$ and $\alpha = 2$. Rates were scaled from SUPERB rates provided by Bhandari et al. (2018).

L-band receiver, multiple operating bands enables e-MERLIN to probe a large spectral parameter space and examine the frequency-dependent behaviour of FRBs[8].

The ability to export high time-resolution, dual circular polarisation voltage data Nyquist sampled at 128 MHz[9] enables fine temporal analysis (on timescales as short as 8 ns) and processing (or reprocessing) of events to form data products with fine spectral resolutions (see filterbanking, Section 5.2.1). This potentially allows the polarisation properties of any detected events to be studied which yield insight into the nature of FRB emission mechanisms or local environments (see, e.g., Michilli et al. 2018).

The telescope's multiple dishes may also be used to expand the FoV available to the array. The HPBW angular resolution of a single 25 m dish at L-band is ∼35 arcmin, corresponding to a FoV of ∼ 0.27 square degrees. Assuming a six-dish fly's-eye configuration (see Section 3.2.1), the array's potential FoV would increase sixfold. This configuration would provide lower sensitivity and worse localisation. However, e-MERLIN's ability to export data separately to regular observations also provides the opportunity to operate a commensal fast radio transient detection program which has no impact on the facility's regular observing routine.

During regular observations when e-MERLIN's dishes are focused on the same area of sky, the opportunity to perform incoherent beamforming allows a FoV equivalent to that of a single dish to be searched for fast transients with improved sensitivity while reducing the significance of RFI at individual telescopes. By use of anti-coincidence techniques (see, e.g., Wayth et al. 2011) to discard terrestrial RFI local to each dish, a lower threshold S/N for detected bursts could be applied, potentially increasing the number of faint candidates detected. Additionally, further RFI mitigation techniques could be used to reject data from dishes with heavy RFI contamination, decreasing false detection rates further[10].

Most importantly, when used together, the array's maximum 217 km baseline provides e-MERLIN with sub-arcsecond localisation potential via coherent beamforming or imaging. An angular resolution of ∼150" at L-band is adequate for host galaxy association of extragalactic transients (Chatterjee et al., 2017), and precise enough to identify specific regions within galaxies, as was demonstrated during the follow-up effort for FRB 150418 (Keane et al., 2016; Bassa et al., 2016). By storing voltage data

---

[8]Recent FRB detections have begun to show frequency-dependent intensity fluctuations. See, e.g., FRB 151230 Bhandari et al. 2018 for examples.

[9]See Section 4.2.4 for a full explanation of e-MERLIN's temporal resolution capability.

[10]See Section 5.2.3 for a practical implementation with LOFT-e.

around the time of the events themselves, this localisation could be achieved retroactively for any individual transient detected, without the need for repeat events or gridding. A review of e-MERLIN's angular resolution, sensitivity and FoV for the various telescope configurations which were described in Chapter 3 are provided in Table 4.2.

By calculating the associated fluence thresholds

$$F_{\text{min,arr}} = S_{\text{min,arr}} \times \tau, \tag{4.1}$$

for an array with threshold sensitivity $S_{\text{min,arr}}$ for events of width τ, and by scaling recent SUPERB FRB rates[11] to e-MERLIN's sensitivity and FoV following Connor et al. (2016) [12]:

$$R_{\text{e-MERLIN}} = R_{\text{SUPERB}} \times \left(\frac{F_{\text{min,e-MERLIN}}}{F_{\text{min,SUPERB}}} \right)^{-\alpha} \times \left(\frac{\Omega_{\text{e-MERLIN}}}{\Omega_{\text{SUPERB}}} \right), \tag{4.2}$$

where $\Omega_{\text{min,arr}}$ denotes the FoV of the instrument and α is the flux intensity distribution index, the expected FRB detection rate for e-MERLIN, $R_{\text{e-MERLIN}}$ may be obtained.

For a six-dish incoherent beam, the expected event rate at S/N = 10 for $\tau = 1$ ms events for e-MERLIN is ~ 0.002 FRBs day^{-1} at L-band. This calculation assumes a power law flux intensity distribution ($N(F) \propto F^{-\alpha}$ with $\alpha = 3/2$ where N is the number of detections) for FRBs, which is expected for homogeneously distributed sources throughout a Euclidean Universe and is consistent with the most recent SUPERB observations ($\alpha = 2.2^{+0.6}_{-1.2}$, Bhandari et al. 2018). The expected number of events detected by a six-dish e-MERLIN incoherent beam as a function of observing time for different SNRs is shown in Figure 4.2.

The development of a high time-resolution backend for e-MERLIN opens up a new research parameter space previously unavailable to the facility. Although the expected e-MERLIN FRB detection rate is low, by operating commensally to regular observations a fast transient backend could exploit all on-sky time available to the observatory while increasing sensitivity to bursts. By doing so, it may potentially contribute a small number of FRBs to the known population every few years, while maximising scientific output via sub-arcsecond localisation, fine temporal and spectral resolution data products, and access to dual polarisations. It may likewise contribute to the study of Galactic fast transients via, e.g., pulsar timing or RRAT localisation[13]. These goals

---

[11]$R_{\text{SUPERB}} \simeq 1.7^{+1.5}_{-0.9} \times 10^3$ FRBs sky^{-1} day^{-1} (Bhandari et al., 2018).

[12]Connor et al. (2016) used 700-900 MHz observations and assumed a flat spectral index.

[13]See Chapter 5 for more on RRAT localisation.
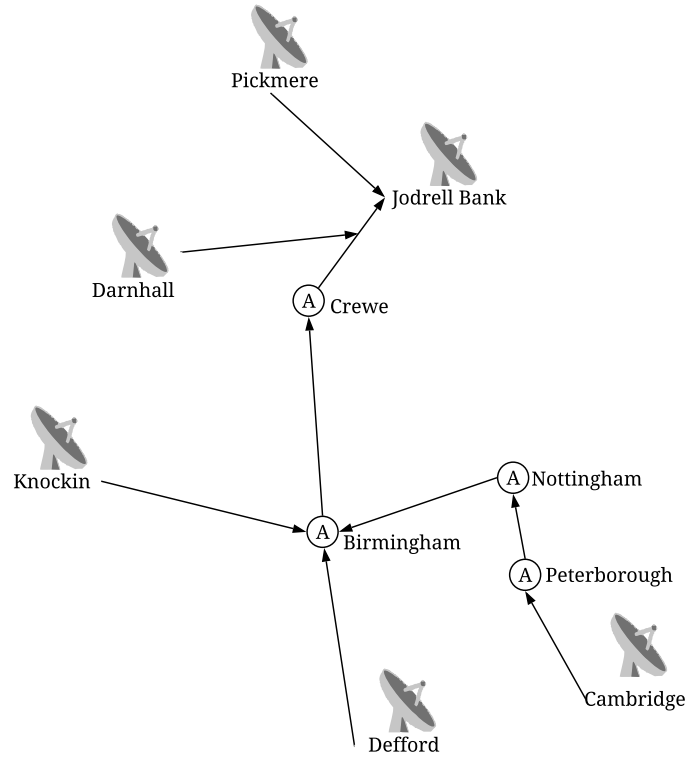
Figure 4.3: The signal chain between individual e-MERLIN telescopes and JBO. Data is digitised by Digital INTerface networks (DINTs) at the telescope sites, before being transmitted. Amplification points (A) exist at Peterborough, Nottingham, Birmingham, Crewe, and JBO. Figure style modelled after Spencer (2009). Correct amplification point schematic from McCool et al. (2006).

have driven the development of the LOFT-e backend, primarily by the author of this thesis, C. R. H. Walker, and fellow PhD student T. W. Scragg.

4.2 The e-MERLIN signal chain

LOFT-e has been developed around pre-existing e-MERLIN hardware and software. This places constraints on the nature of the data available to the backend. Data acquisition by LOFT-e takes place late in the standard e-MERLIN signal chain, by which time it has undergone several obligatory stages of processing including amplification, digitisation, resampling and filterbanking, and multiple stages of quantisation. These stages influence the quality of the data recieved by LOFT-e and the format of its output, dividing its 512 MHz bandwidth into 8 sub-bands of dual polarisation, 2-bit, 128 MHz sampled voltage data.

It has been important to understand these stages of data transformation over the course of LOFT-e's development. The transformations undergone affect the bandwidth, sampling time, and sensitivity of observations. Even when implemented optimally, the conversion of analogue data to a digital format alone reduces sensitivity (however these losses can be small, see Section 4.2.1.1). When implemented suboptimally, e-MERLIN's own RFI mitigation techniques may negatively impact the Gaussianity of data, as will be discussed in Section 4.2.4. Certain parts of the e-MERLIN signal chain may actively introduce RFI to observations, as will be discussed in Section 4.5. These effects have been observed over LOFT-e's development cycle, and observing strategies and RFI mitigation techniques have been implemented accordingly. Additionally, throughout the project feedback has been provided to e-MERLIN staff consistently to allow informed development of e-MERLIN's own backends.

Therefore in this section the major stages of data transformation undergone enroute to LOFT-e are introduced. Their potential impact on LOFT-e's sensitivity is discussed. Where applicable, changes which have been made to e-MERLIN's backend are noted, and places where further investigation is necessary are identified.

4.2.1   From telescope to correlator

The path between e-MERLIN dishes and its correlator is shown in Figure 4.3. Analogue data from each individual telescope receiver is primarily sampled by an analogue-to-digital converter known as a Digital INTerface (DINT) at the telescope site. The resulting right and left-hand circularly-polarised digital signals, multiplexed onto 3×10 Gb s^{-1} light paths per antenna, are transferred across a network of dark fiber-optic cables. The total data rate per telescope is ~ 30 Gb s^{-1} (Garrington et al., 2004). At this stage in the signal chain, data are 8-bit voltage data in a proprietary format.

En-route to Jodrell Bank Observatory, data from different antennae may pass through several amplification points consisting of Erbium Doped Fibre Amplifiers (EDFAs)[14]. Telescopes stationed furthest from the correlator (at Knockin, Defford and Cambridge) are directed to an amplification point in Birmingham where they are redistributed to a single fibre. Cambridge has two stages of EDFA amplification prior to Birmingham, at Peterborough and Nottingham. There are subsequent amplification points at Crewe, and at Jodrell Bank Observatory (JBO). At Birmingham and Crewe, the light paths for Knockin, Defford and Cambridge data are combined onto a single fibre and transmitted

---

[14]https://www.rp-photonics.com/erbium_doped_fiber_amplifiers.html   (Accessed: 25/09/2018).
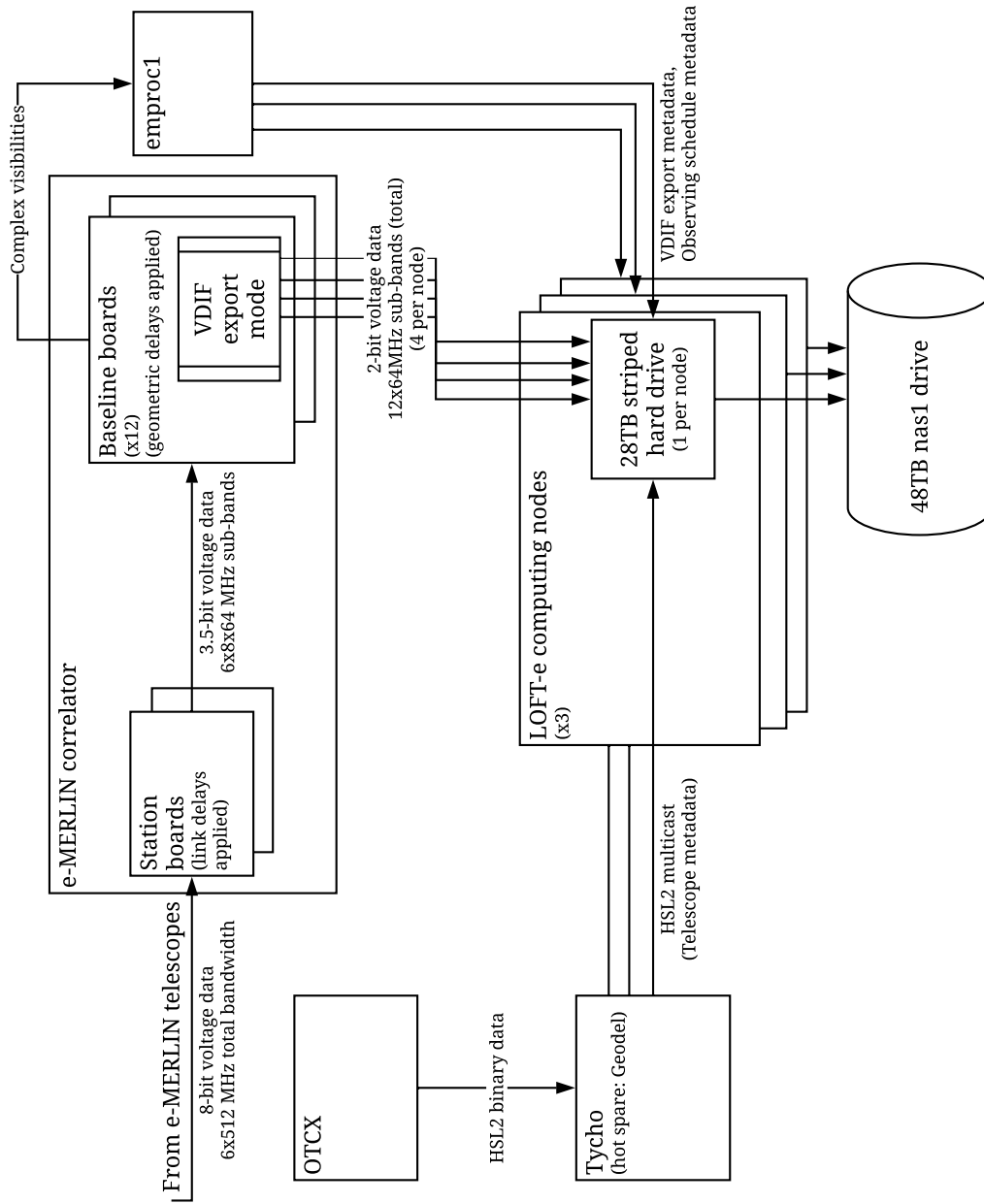
Figure 4.4: System flow schematic for e-MERLIN data and metadata as received by LOFT-e. Upon reaching Jodrell Bank Observatory, 512 MHz bandwidth of voltage data enters the e-MERLIN correlator. Link delays are applied at the station boards, data are divided into 8×64 MHz sub-bands and requantised to 3.5-bit. Data then enters baseline boards where geometric delays are applied and correlation is performed. VDIF export mode forks and requantises voltage data to 2 bits. This is extracted by LOFT-e. Telescope, VDIF export and observing schedule metadata are also acquired from various machines.

via nine separate EDFA colours. Data from the closer telescopes (Pickmere, Darnhall, Mark II) are transmitted to JBO directly.

Upon reaching the observatory, the 8-bit voltage data enters the e-MERLIN Correlator. The correlator is a modified version of the Wideband Interferometric Digital ARchitecture (WIDAR) corellator developed for the Expanded Very Large Array (EVLA)[15]. Detailed information and references regarding the e-MERLIN correlator may be found on the Jodrell Bank Observatory Operations webpage[16]. For the purposes of this work it is necessary to know that the correlator contains *station boards* and *baseline boards*. Data transformation stages within these boards are discussed below. The signal path for data within the correlator is shown in Figure 4.4.

To account for thermal length variations of the optical fibres connecting telescopes to the correlator, thus ensuring coherence across the array, e-MERLIN utilises a 'Timing Link' system. Pulsed signals of frequency 1486.3 MHz, locked to a Sigma Tau hydrogen maser source, are transmitted to a directly modulated laser of central wavelength 1550.31 nm at JBO. These modulated optical signals are then transmitted over e-MERLIN's fiber-optic network (see Figure 4.3) and recieved at each telescope. The recieved signal is used to a lock quartz oscillator at each site to a 499.9 MHz signal which is then output to a further laser of central wavelength 1550.85 nm and transmitted back to JBO. At JBO, round-trip phase differences between the initial and returned signal modulations are calculated for each telescope, and corrections corresponding to half of these delays are applied to incoming data. In order to distinguish between timing signals, the pulsed modulations vary in pulse length between telescopes (McCool et al., 2008; Thomasson, 1986).

4.2.1.1   Sampling and digitisation

During the process of digitisation, a signal is sampled periodically and quantised into specific values. The *Nyquist criterion* specifies that to accurately reproduce the waveform of a signal incoming to a radio telescope, real data must be sampled at least at the *Nyquist Frequency* $\nu_{\mathrm{Nyq}} = 2 \times \Delta\nu$ where $\Delta\nu$ is twice the observational bandwidth (Hagen & Farley, 1973). Although e-MERLIN's original bandwidth is 512 MHz, its data are split into multiple 64 MHz wide frequency sub-bands in the correlator (see Section 4.2.2). Thus data recieved by LOFT-e is sampled once every $1/(128 \times 10^6)$

---

[15]For more details, see: `http://www.aoc.nrao.edu/widar/docs/Memos/memos/nrc_evla_memo001.pdf` (Accessed: 25/09/2018).

[16]`redmine.jb.man.ac.uk` (Accessed: 25/09/2018)
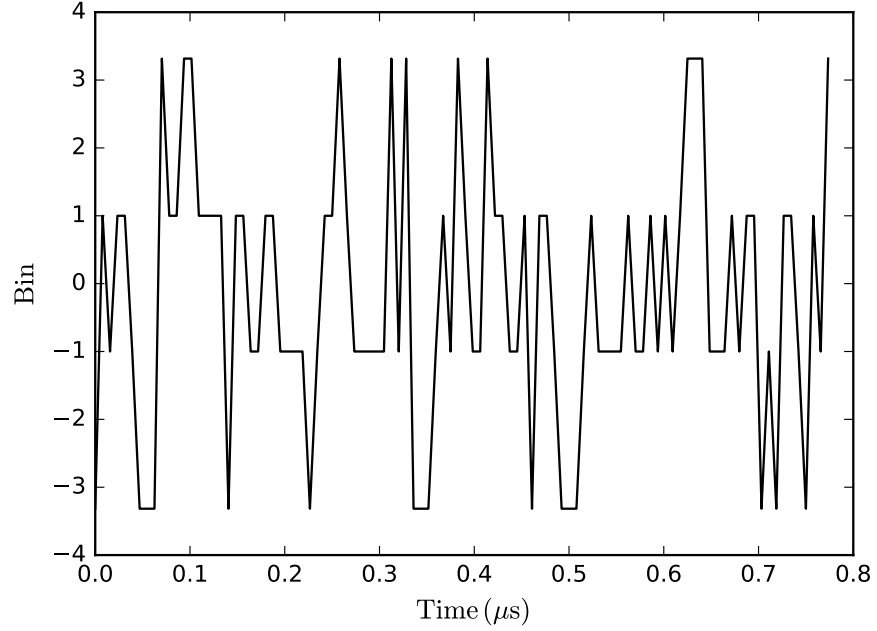
Figure 4.5: Time into an observation plotted against bin number, showing a sample of the 2-bit digitised signal aquired by LOFT-e from e-MERLIN's Cambridge Telescope during L-band observation of PSR B1933+16. Data is sampled at 128 MHz.

Number of bits, n	Number of bins, 2^n	Digitisation factor, β
1	2	0.636620
2	4	0.881154
4	16	0.988457
8	256	0.999912

Table 4.3: Digitisation efficiency factor β for different digitisers. An n-bit digitiser may sort data into a potential 2^n different bins. More bits assigned to individual time samples equates to more bins for data representation, and less SNR degradation occurs. Values are obtained from Thompson et al. (2007).

seconds.

During digitisation, analogue voltage data comprising an infinite number of potential values is approximated as digital values. The possible digital values, or bins, that the data may be approximated by is discrete, and is defined by the number of bits of information assigned to a single time sample. A larger number of bits allows for a larger dynamic range comprised of more bins. This will more accurately represent the true data at the expense of larger data rates. Smaller numbers of bits result in lower data rates at the cost of less accurate representation of the original data and higher SNR loss (Burke & Graham-Smith, 2002; Bowers et al., 1973; Cooper, 1970). All digitisation results in some SNR loss, represented by the digitisation factor β in the radiometer equation (see Chapter 3), however this is nearly negligible for >8-bit data. Ideal SNR losses for different digitisers are shown in Table 4.3. An example of the 2-bit digitally sampled signal which is captured by LOFT-e is shown in Figure 4.5.

An n-bit digitiser assigns n bits of information to each time sample, affording a potential 2^n bins for data representation. Typically, a 2-bit digitised time sample will be represented by 'sign' value (indicating positive or negative) and a 'magnitude' value. For Gaussian data[17] with zero mean, typically time samples with values $< 1\sigma$ from the mean will be sorted into the central two bins. All remaining data will fall into the outer bins unless clipping or some other form of RFI mitigation is performed (see Chapter 5).

Three quantisation stages occur during e-MERLIN's signal chain: from analogue to 8-bit at the DINTs, from 8-bit to 4-bit at the station boards (see Section 4.2.2) and finally from 4-bit to 2-bit (see Section 4.2.4). The final output as aquired by LOFT-e is of VLBI Data Interface Format (VDIF). This data format is mean subtracted, and within e-MERLIN time samples are assigned values of -2, -1, 1, and $2^{18,19}$. Digitisation-induced sensitivity loss for LOFT-e is a combination of the 2-bit, 4-bit, and 8-bit β values provided in Table 4.3.

---

[17]Gaussianity is a fundamental assumption of the radiometer equation (see Chapter 3).

[18]https://vlbi.org/vdif/docs/VDIF_specification_Release_1.1.1.pdf   (Accessed: 25/09/2018)

[19]Bins may be assigned any value. If assigned such that the inner bins lie one standard deviation (σ) from a mean-subtracted Gaussian with $\sigma = 1$, then the optimal representation of the true data is -3.316505, -1, 1, and 3.316505.
During analysis of VDIF data, LOFT-e makes use of the Python package, baseband (https://github.com/mhvk/baseband), which reassigns these values optimally (https://baseband.readthedocs.io/en/stable/api/baseband.vlbi_base.encoding.OPTIMAL_2BIT_HIGH.html) (Accessed: 25/09/2018).

4.2.2 The correlator station boards

Upon reaching the e-MERLIN correlator, dual-polarisation, 512 MHz bandwidth voltage data for each telescope arrives in a proprietary 8-bit format and enters the correlator. Within the correlator data are first directed through station boards[20]..

Within the station boards, *link delays* are applied for each telescope. The link delay accounts for light travel time through the fiber-optic cables between the telescopes and correlator. This delay is not constant with time as the fibers will expand and contract with temperature fluctuations. Additionally, within the station boards data are requantised from 8-bit to 4-bit data. During this requantisation stage, the data are mapped to 15 bins[21]. This 15-bin data are colloquially referred to as 3.5-bit data.

Additionally, during this stage in the signal chain Fourier transforming and finite impulse response (FIR) filtering (polyphase filterbanking) are performed on the data, dividing e-MERLIN's continuous 512 MHz bandwidth into eight 64 MHz wide subbands, sampled at $\nu_{Nyq} = 128$ MHz.[22]

4.2.3 The correlator baseline boards

After exiting the station boards, the 3.5-bit, 8×64 MHz, dual-polarisation sub-bands enter the correlator's baseline boards[23]. Within the baseline boards, *geometric delays* are applied for each telescope. The geometric delays account for the differing positions of e-MERLIN's dishes on Earth relative to source positions in the sky, effectively shifting the locations of each dish to a nominal center of the array. In addition to geometry and Earth rotation effects, the geometric delay also takes into account more complex, smaller effects (e.g a general relativistic time delay due to differences in gravitational fields at different heights above the Earth's surface).

During standard interferometric operations correlation (multiplication and averaging) is performed on the 3.5-bit data at the baseline boards, creating complex visibilities. Following correlation, the visibilities are sent elsewhere[24] for further processing[25]

---

[20]The e-MERLIN correlator accepts data with a central frequency channel, so during 8-bit digitisation data are mapped to only 255 bin

[21]Again, this 4-bit data has 15 bins instead of 16 bins to provide a central frequency channel.

[22]The eight L-band sub-band center frequencies are: 1286.4, 1350.4, 1414.4, 1478.4, 1542.4, 1606.4, 1670.4 and 1734.4 MHz.

[23]For more information on the baseline boards, see: `http://www.aoc.nrao.edu/widar/docs/BaselineBoard/rfs/A25080N0001_BaselineBoardUserManual_V1.1_Feb10-12.pdf` (Accessed: 25/09/2018).

[24]Initieally emproc1, then periodically moved to magnetic tapes and emvault2 by schedulers.

[25]E.g. more time averaging.

and storage.

The integration timescale for visibilities is specified at the time of correlation. The theoretical minimum integration time for e-MERLIN visibilities is 250 ms, however hardware and software constraints currently limit output to a minimum integration period of 1 s. As discussed in Chapter 3, this timescale is long enough to negatively impact fast transient detection in the image plane. Shorter timescale visibilities could be achieved through software correlation. The geometric delay model currently applied to e-MERLIN data are non-standard. Investigations are currently taking place to determine whether it could be replaced by the standard VLBI delay model CALC11 (Gordon et al., 2016), which is commonly used in software correlators such as SFXC (Keimpema et al., 2015) and DIFX (Deller et al., 2011).

The number of functioning baseline boards determines the total amount of data which e-MERLIN may output to LOFT-e. In total, the correlator has 16 baseline boards. As of writing 12 are functional, allowing 12×64 MHz dual-polarisation sub-bands to be distributed to LOFT-e for further processing.

4.2.4   VDIF export mode

LOFT-e exploits e-MERLIN's VDIF export mode in order to aquire voltage data from the telescopes. This operating mode exists to allow e-MERLIN integration with the European VLBI Network (EVN) (Spencer, 2009). The mode is currently still under commissioning and ~ 36 hours of observing time is allocated to VLBI every six months.

Within the baseline boards, VDIF export mode forks e-MERLIN's 3.5-bit, 128 MHz sampled voltage data, requantises it to 2-bit, and exports it. Each e-MERLIN baseline board exports a dual polarisation 64 MHz wide sub-band. By streaming the data from a baseline board to an individual port on a LOFT-e computing node, the LOFT-e back-end acquires 2-bit voltage data sampled at 128 MHz for fast transient analysis. Each 64 MHz wide chunk may be assigned to an e-MERLIN telescope and frequency sub-band chosen by the LOFT-e user. At the time of writing, 12 baseline boards are functional, allowing a total bandwidth of 768 MHz split across telescopes and frequency sub-bands to be streamed to LOFT-e nodes.

During VLBI observations geometric delays are *not* applied to data by the e-MERLIN correlator. This results in the creation of no useful e-MERLIN visibilities, and individual telescope beams formed with LOFT-e data processing pipelines will not
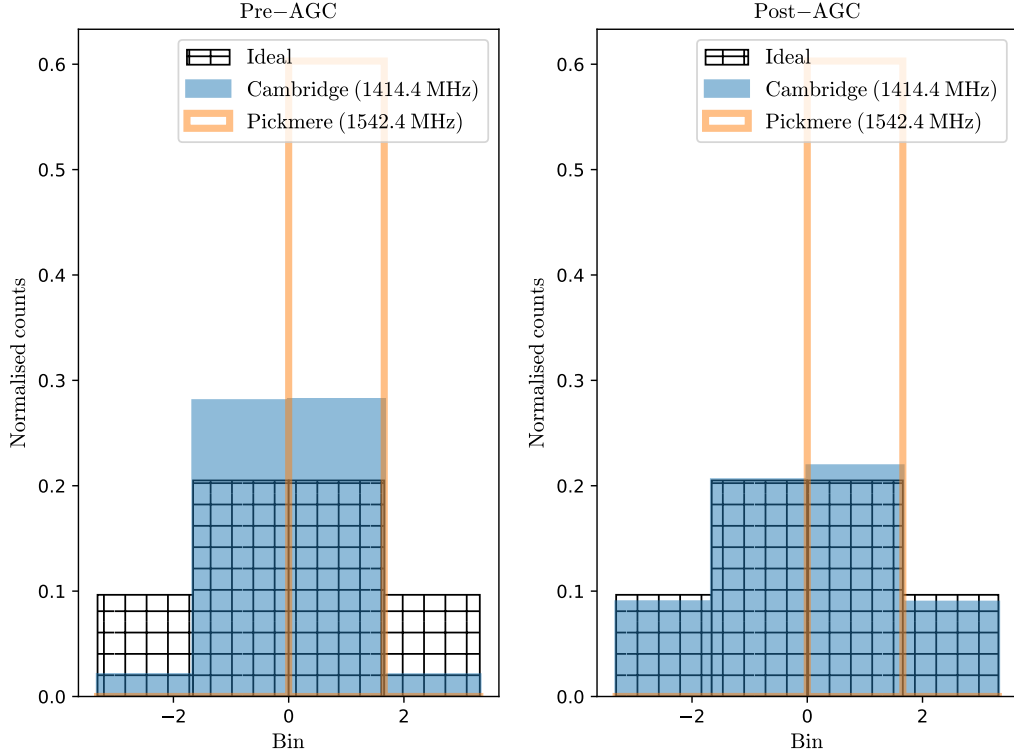
Figure 4.6: Normalised histograms of sample counts against bin number for 2-bit VDIF data streamed simultaneously to LOFT-e from e-MERLIN's Cambridge (solid blue) and Pickmere (orange outline) telescopes. **Left:** distribution prior to running e-MERLIN's AGC command. **Right:** distribution post-AGC command. A cross-hatch of an ideal Gaussian distribution is provided for reference. Note that Pickmere data are saturated and so reside only in a single bin.

occur at the same time - this delay would need accounting for during incoherent beam-forming. However during non-VLBI observations, geometric delays are applied, and exported VDIF data are contemporaneous between telescopes.

4.2.4.1 The auto gain control system

During VDIF export mode's requantisation of data from 3.5-bit to 2-bit, a scaling factor is applied to each data stream. This factor is calculated using e-MERLIN's auto gain control (AGC) system.

This scaling factor influences which of the four bins available to 2-bit data a time sample will be assigned to during the requantisation process which was overviewed in Section 4.2.1.1. Sub-optimal mapping of data during this requantisation stage may result in a collection of samples which deviates from a Gaussian distribution. All time samples may be mapped to the inner two bins, effectively resulting in single-bit data

which will reduce S/Ns of observations. A significantly extreme scaling factor may saturate the data, destroying all information it contains by mapping all time samples into a single bin (see, e.g., the Pickmere telescope data distribution in Figure 4.6).

An example of Cambridge telescope data distributions prior to and post AGC triggering is shown in Figure 4.6. Before the AGC is triggered, Cambridge data deviates from a Gaussian distribution. Post-AGC, Cambridge data displays a more Gaussian profile[26].

The AGC ideally exists to automatically adjust the scaling factor, accounting for sporadic RFI changes and variations in background sky temperature as a source being observed transits across the sky. However, VDIF export mode is still being commissioned and e-MERLIN's AGC is still being observed to behave sub-optimally. The software has gone through several iterations during LOFT-e's development cycle. These iterations have been implemented into e-MERLIN's backend by observatory staff, and have been informed by LOFT-e testing. At the time of writing, the AGC is applied during LOFT-e observations manually by issuing a single command to e-MERLIN's correlator. The scaling factors are subsequently calculated individually for the data of each e-MERLIN telescope, frequency sub-band, and polarisation. Previous AGC tests, and how best to apply the most recent iteration of e-MERLIN's AGC is discussed in Chapter 6.

4.3   The e-MERLIN metadata chain

In order for LOFT-e to extract meaningful results from e-MERLIN, data aquisition alone is not adequate. Additional metadata regarding telescope configurations, observing schedules, and data export configurations are all necessary to understand and utilise the data being streamed from the correlator's baseline boards to LOFT-e. Examples of necessary information include:

1. Mapping structure of baseline boards to LOFT-e's computing nodes. This is necessary to identify the sources of individual data streams, enabling the behaviour of each to be analysed individually. This may be useful for, e.g, identifying and excluding telescopes during beamforming which are off source or undergoing maintenance.

---

[26]Note: Post-AGC triggering, Pickmere's distribution is not affected. During this observation Pickmere was undergoing routine maintenance: its non-Gaussianity was occuring earlier in the e-MERLIN signal chain, and the AGC was not expected to improve the underlying distribution.

2. Current telescope technical configurations such as the frequency band (i.e. L-band, C-band, K-band) being observed. This enables correct analysis of frequency-dependent behaviour of transients (for example dispersion). By discriminating between frequency sub-bands, analysis of RFI severity at different dishes as a function of sky location or time of day may also be performed[27].

3. Routine logging of e-MERLIN dish antenna pointings on the sky during data recording. This enables localisation in case of a commensal FRB detection, and ensures telescopes are on-source during dedicated observations of pulsars (or other targets).

The three main sets of metadata necessary for LOFT-e to function are: *telescope metadata*, which describes the status of each telescope at any time; *VDIF export metadata*, which describes the current mapping of baseline boards to LOFT-e; and *observing schedule metadata*, which describes the scheduled sources to be observed by the array. A flow diagram of all data and metadata received by LOFT-e may be found in Figure 4.4. The three sets of metadata are discussed in detail below.

4.3.1   Telescope metadata

During any dedicated observation over a long period, e-MERLIN telescopes must first slew to the target source and then track its motion across the sky. Information regarding whether individual dishes are functional, on source, and tracking correctly at any point in time is contained within metadata colloquially known as High Speed Link 2 (HSL2) messages.

HSL2 messages are initially broadcast by e-MERLIN's main control computer, OTCX, in a binary format over Jodrell Bank Observatory's restricted access telescope control network. The messages are passed through a chain of programs on an intermediary computer where they are restructured and rebroadcast as multicast messages on the observatory's house network.

During HSL2 broadcasting, eleven 200-byte different messages are broadcast sequentially, one per second, over an eleven-second cycle. The first two messages contain clock and meteorology information respectively. The next nine consecutive messages contain metadata regarding different JBO telescopes (including instruments which are not part of e-MERLIN). This metadata includes:

---

[27]As has been done by PhD student T. Scragg to inform sub-band choices during observations.

1. Telescope name,

2. Telescope motor status: may be 'on' or 'off'.

3. Current target source name.

4. Current and demanded azimuth and altitude coordinates.

The final format for decoding HSL2 messages as they are received by LOFT-e is described in detail in MERLIN Document 28[28] LOFT-e's computing nodes subscribe to the multicast group and receive the messages over LOFT-e's ethernet 2 (eth02) link. Messages are routinely captured and logged by the LOFT-e data capture pipeline (see Chapter 5) once per minute.

During dedicated e-MERLIN observations of repeating sources such as pulsars, it is often useful to observe and store data over long periods of time, capturing many single pulses to enable folding (see Chapter 5) and the associated increase in sensitivity resulting from combining single pulses. This is useful for scientific purposes such as pulsar timing. It is also useful for testing of LOFT-e's pipelines, for example to quantify the effectiveness of incoherent beamforming or RFI mitigation techniques (see Chapters 5 and 6).

If telescopes are not tracking a requested source during a targeted high time-resolution observation, the observation is not being performed correctly. In such a case, e-MERLIN's large data rates (which may reach up to 768 MB/s) and LOFT-e's limited storage capacity necessitate that data should ideally not be recorded. However during observations of other sources (e.g. observations for legacy projects or accepted e-MERLIN proposal holders), LOFT-e may still operate commensally, forking and capturing data and searching it for new radio transients. When complete, LOFT-e's commensal mode will process data in real-time and only a few seconds of voltage data around candidate fast radio transient events will be stored for further processing and localisation[29]. During both targeted and commensal observations telescope pointing information is required to localise events on sky. Therefore HSL2 metadata are vital for both LOFT-e's commensal and non-commensal modes of operation.

---

[28]Obtained from C. Jordan via priv. com.

[29]A full commensal mode is not yet fully operational.

4.3.2   VDIF export metadata

The voltage data streamed from e-MERLIN's baseline boards to LOFT-e contains no identifying information which can be used to match a data stream to its respective telescope or sub-band centre frequency. Without this information, a given telescope may not be discarded during beamforming if off-source or under maintainance during observations. Additionally, the frequency-dependent nature of dispersion means that a fast transient event will be observed at slightly different times in different e-MERLIN sub-bands. This must be accounted for during beamforming and sub-band combination. The information which contextualises data streams and maps them from e-MERLIN baseline boards to LOFT-e's computing nodes will hereafter be named VDIF export metadata.

VDIF export metadata are broadcast in JavaScript Object Notation (JSON) data format by e-MERLIN's `emproc1` machine. It is queried directly by LOFT-e[30] using tools developed for the LOFT-e data capture pipeline[31]. The metadata takes the form of JSON data dictionaries and contain the following identifying information:

1. Board: Baseline board identification number and LOFT-e port number that data are being streamed to. Takes the form of a five-digit integer.[32]
2. IP: Internet protocol address of the data stream's target LOFT-e computing node.
3. Telescope: Two-character code denoting the data stream's telescope of origin.
4. effectiveLO: Current effective telescope local oscillator frequency (in Hz).
5. CentreFreq: Centre frequency (in Hz) of the stream's sub-band after subtraction of effectiveLO.
6. Bandwidth: Bandwidth of the data stream (in Hz).
7. Pol: Circular polarisation of data stream. Takes the form left/right : true/false.
8. requantizerGain: Scaling factor calculated by the AGC (see Section 4.2.4.1) which is being used to rescale data.

VDIF export metadata are recorded alongside e-MERLIN telescope data by the LOFT-e data capture pipeline and used by LOFT-e data analysis pipelines during data processing. It is always stored as a `PortConfig.txt` file in the same folder as the raw voltage data.

---

[30] `curl http://emproc1.ast.man.ac.uk:8080/correlator/vdif`

[31] `/home/cwalker/loft-e/lofte_receivers/eJSON_utils.py`

[32] Beginning `130XX`.

4.3.3   Observing schedule metadata

To achieve time and frequency resolutions useful for fast radio transient analysis, further filterbanking (see Section 4.2.2) is performed by the LOFT-e data analysis pipeline. Output data products are stored in the format standardised by the pulsar signal processing software, *sigproc*. These products are then combined for e-MERLIN's telescopes during beamforming.

Sigproc filterbanks require header information including bandwidth, center frequency, and spectral resolution of the data. They also require timestamps which identify when data was recorded. Timestamps are necessary during the process of beamforming in order to ensure data to be combined is contemporaneous. Timestamps are provided within the VDIF headers of the recorded voltage data. However, VDIF headers contain no information on the target source which was being observed by e-MERLIN at the time of data capture. Source information is necessary for discrimination between datasets (e.g., of observing targets vs. phase calibrators) in order to perform coherent beamforming or imaging. This information is provided by e-MERLIN's observing schedule metadata.

As with VDIF export metadata, observing schedule metadata is broadcast as JSON messages by `emproc1` and queried directly by LOFT-e[33]. This metadata contains information regarding all e-MERLIN observing jobs which have been scheduled at the time of a query call including:

1. List of target source names.
2. UTC time observing will begin for targets.
3. Telescope drive times and directions when transitioning between targets.
4. Telescope observing band of observation (e.g. L-band, C-band, K-band).

Observing schedule metadata are recorded by the LOFT-e data capture pipeline. During LOFT-e filterbanking, VDIF header times are cross-referenced against the times in this observing schedule in order to populate sigproc filterbank headers with source names and timestamps prior to beamforming.[34] As with VDIF export metadata, observing schedule metadata are stored alongside raw voltage data in its respective directory as a `Schedule.txt` file.

---

[33]`curl http://emproc1.ast.man.ac.uk:8080/ojd/current`
[34]`/share/nas1/LOFT-e/software/analysis_pipeline/make_gpufil_headers_2.py`

4.4 The LOFT-e hardware backend

Upon leaving the baseline boards, e-MERLIN's 2-bit voltage data are streamed to LOFT-e's hardware backend for further processing and fast radio transient analysis (see Chapter 5). Data is streamed via ethernet packets along fibre-optic cables to one of the local area network (LAN) interfaces on each of LOFT-e's computing nodes. Data streams are permanently active[35] which will enable commensal data processing and single-pulse searching when a real-time version of LOFT-e's data processing pipeline (see Chapter 5) is complete.

The LOFT-e hardware backend consists of three computing nodes[36] named `transient`, `compute-0-0`, and `compute-0-1`. As at the time of writing 12 of e-MERLIN's 16 baseline boards are functional it is advisable to distribute these streams evenly across computing nodes to balance data rates. Where possible, sub-bands from the same telescope should also be directed to the same computing node. Assuming a fully-functional six-telescope e-MERLIN array, best practice thus dictates that each node receives 4 data streams, consisting of $2\times$ subbands for $2\times$ telescopes. Non-standard data distributions are possible, however, and may be decided by a LOFT-e user in conjunction with e-MERLIN staff.

As discussed in Sections 4.2.1.1 and 4.2.4, the e-MERLIN correlator's VDIF Export Mode outputs 2-bit voltage data as UDP packets containing VDIF data frames. Each frame contains 32000 2-bit time samples of size 8032 bytes, consisting of a 32-byte header and 8000 bytes of data[37]. These UDP Packets are transmitted at a rate of 4000 packets per polarisation per second[38]. For the standard 12 baseline board setup described above, the resulting data rate is 256 MB s^{-1} per LOFT-e computing node.

Each computing node is equipped with $2\times$ Nvidea Geforce GTX 980 Ti GPUs for data processing[39] (which is detailed in Chapter 5). Each node is additionally equipped with a Nvidea Geforce GT 730 GPU for display purposes.

Raw voltage data and individually processed telescope sub-bands are stored to

---

[35]'Permanently' excludes during time spent in e-VLBI observing mode, as current e-MERLIN infrastructure does not allow the correlator to redirect data to multiple addresses at once.

[36]Intel Core i7-5930K 3.5 GHz CPUs with 6 Cores / 12 Threads and 32 GB RAM.

[37]For a full specification of VDIF data, see: `https://vlbi.org/vdif/docs/VDIF_specification_Release_1.1.1.pdf` (Accessed: 25/09/2018).

[38]Polarisations are interleaved.

[39]Faster than real-time, GPU-based VDIF filterbanking software (developed by M. Malenta) has superceded previous, slower Python-based filterbanking software developed by the author using the `baseband` Python module (`https://github.com/mhvk/baseband`; Accessed: 25/09/2018). RFI mitigation and incoherent beamforming currently still operate slower than real-time. Faster than real-time GPU-based single-pulse searching software is also utilised. See Chapter 5 for further discussion.

28 TB striped hard drives on each computing node. Beamformed data are stored on a 48 TB Network A Storage (NAS) hard drive shared by all three computing nodes. A full list of LOFT-e hardware is presented in Table 4.4.

4.5   RFI present in e-MERLIN data

This section presents e-MERLIN timeseries data which has been successfully captured by LOFT-e. Analysis of recent observations has uncovered short timescale, quasi-periodic modulation in e-MERLIN data which negatively impact S/Ns recovered during test pulsar observations. These modulations are shown for bandpass-subtracted Stokes I data in Figures 4.7 and 4.8.

Each plot shows a one-second dynamic spectra and associated timeseries extracted from the beginning of a 10-minute LOFT-e observation. The RFI behaviour observed in these data occurs consistently throughout the observations. Figure 4.7 shows simultaneously recorded data compared for three Cambridge sub-bands at different center frequencies, and one Knockin sub-band. Figure 4.8 shows simultaneously recorded data for one sub-band of each of the Darnhall, Pickmere, Mark II and Defford telescopes.

Both the dynamic spectra and timeseries plots display regularly occuring RFI occuring in short bursts. The RFI behaviour is broad-band, spanning the entire 64 MHz of bandwidth in all but the Darnhall telescopes. The associated timeseries plots show that the RFI ocurring in each sub-band displays unique structure. Cambridge, Defford, Darnhall, Pickmere and Mark II RFI appears quasi-periodic. Additionally, RFI for each Cambridge sub-band appears to occur contemporaneously.

Figure 4.9 presents FFT frequency analysis of each timeseries. As shown by the top left subplot, all Cambridge sub-bands display a periodicity of ∼28 Hz with many harmonics. The Pickmere telescope also displays a periodicity at ∼25 Hz with many harmonics. Other dishes show less clear complex structure.

Figures 4.10, 4.11 and 4.12 show the behaviour of the individual polarisations for each data set. In these plots no bandpass has been subtracted. Figure 4.10 compares all Cambridge sub-bands, Figures 4.11 and 4.12 compare remaining telescopes. The structural characteristics of the RFI appear similar between some telescopes and sub-bands (see, e.g., Cambridge 1414.4 MHz and 1542.4 MHz, and Mark II 1414.4 MHz pol 1 data) however the structure of modulations in other data sets is more complex.

More investigation is necessary to determine and eliminate the source of this RFI.
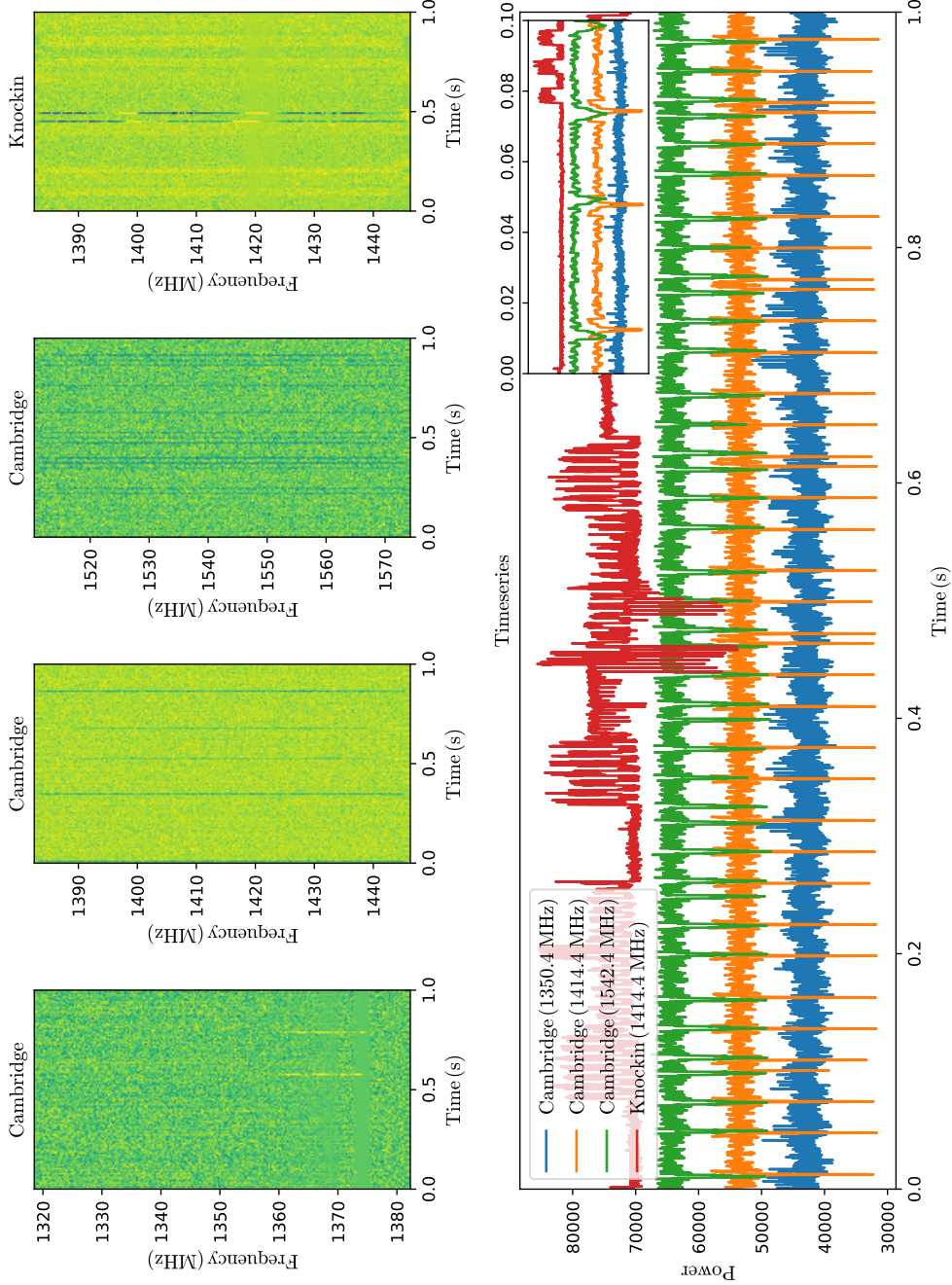
Figure 4.7: **Top:** One-second, 64-MHz-wide dynamic spectra of observing frequency against time for data acquired simultaneously by LOFT-e from multiple e-MERLIN telescopes and frequency sub-bands during an observation of PSR B0329+54. Spectra are Stokes I and bandpass-subtracted. Spectra have 64 μs temporal resolution and 0.25 MHz spectral resolution. Colour indicates relative intensity. **Bottom:** 'frequency-scrunched' timeseries for each dynamic spectrum, offset by a constant value **Inset:** close-up of the first 0.1 s of each timeseries.
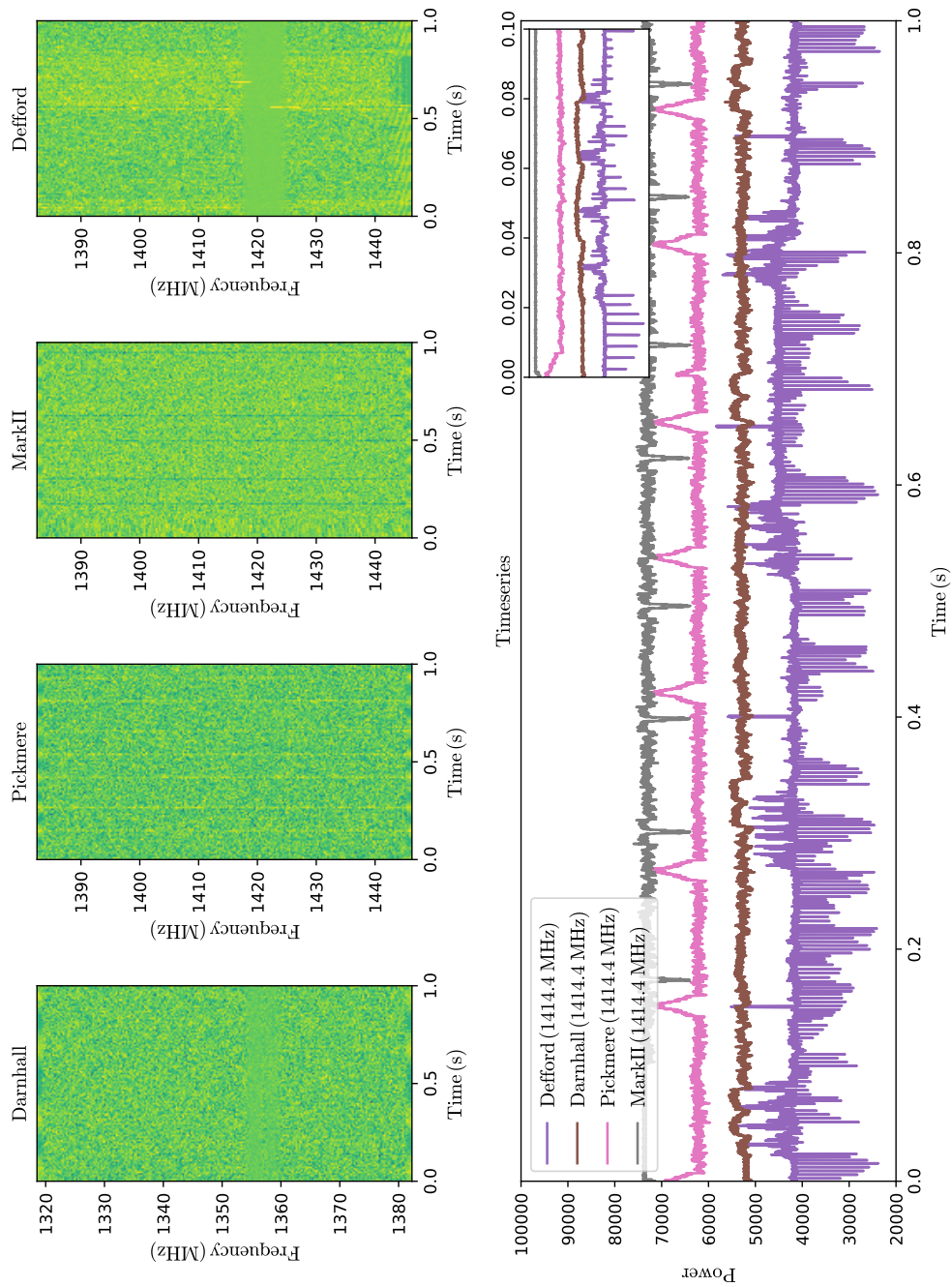
Figure 4.8: As Figure 4.7, for one sub-band center frequency for each of the remaining e-MERLIN telescopes. All data was acquired simultaneously.
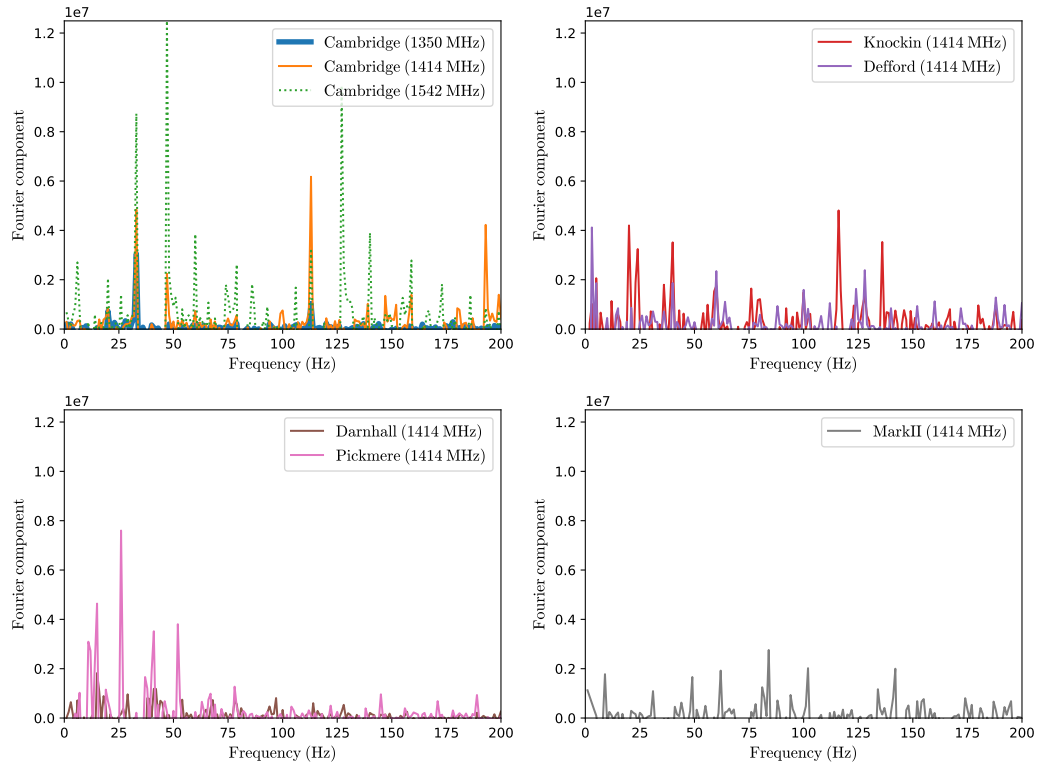
Figure 4.9: Power spectrum plots of Fourier component against observing frequency for each e-MERLIN telescope 1 s timeseries from Figures 4.7 and 4.8, revealing periodicities at ~ 25 Hz for each telescope. The periodicity is identical (28 Hz) between Cambridge sub-bands, but differs between individual dishes.

Different periodicities between telescopes indicate potentially different origins. The identical periodicity across Cambridge sub-bands indicates the RFI occurs prior to division of data into multiple sub-bands by the correlator station boards. All e-MERLIN telescope noise diodes operate at the same switching frequency of 20 Hz. Inconsistencies between this period and the observed periodicities for each telescope and sub-band rule out a noise diode origin for these modulations.

While no origin has yet been determined for this RFI, mitigation techniques have proven effective in reducing its impact on S/Ns recovered during pulsar observations. These techniques are discussed in Chapter 5.

4.6   Future improvements to LOFT-e signal chains

The method used by LOFT-e to receive telescope metadata (see Section 4.3.1) is now outdated. It contains unneccessary points of failure including intermediary computers and metadata rebroadcasting steps. Additionally, binary HSL2 messages may be updated in the future to accomodate extra status information. Contact with observatory staff must therefore be maintained to ensure any HSL2 formatting changes are accounted for.

LOFT-e has direct access to the restricted telescope control network and e-MERLIN's main control computer, `OTCX`. It may thus theoretically query telescope status details from `OTCX` directly, removing redundancies in the metadata signal chain.

Future upgrades to LOFT-e could use existing HSL2 receiver code as a basis for decoding binary `OTCX` messages directly, as is done by Jodrell Bank's pulsar machine, `Ector`. Alternatively, a LOFT-e ugrade could utilise Python-based software which accesses separate XML-based telescope messages to extract necessary telescope metadata. This software is currently being commisioned by e-MERLIN staff.
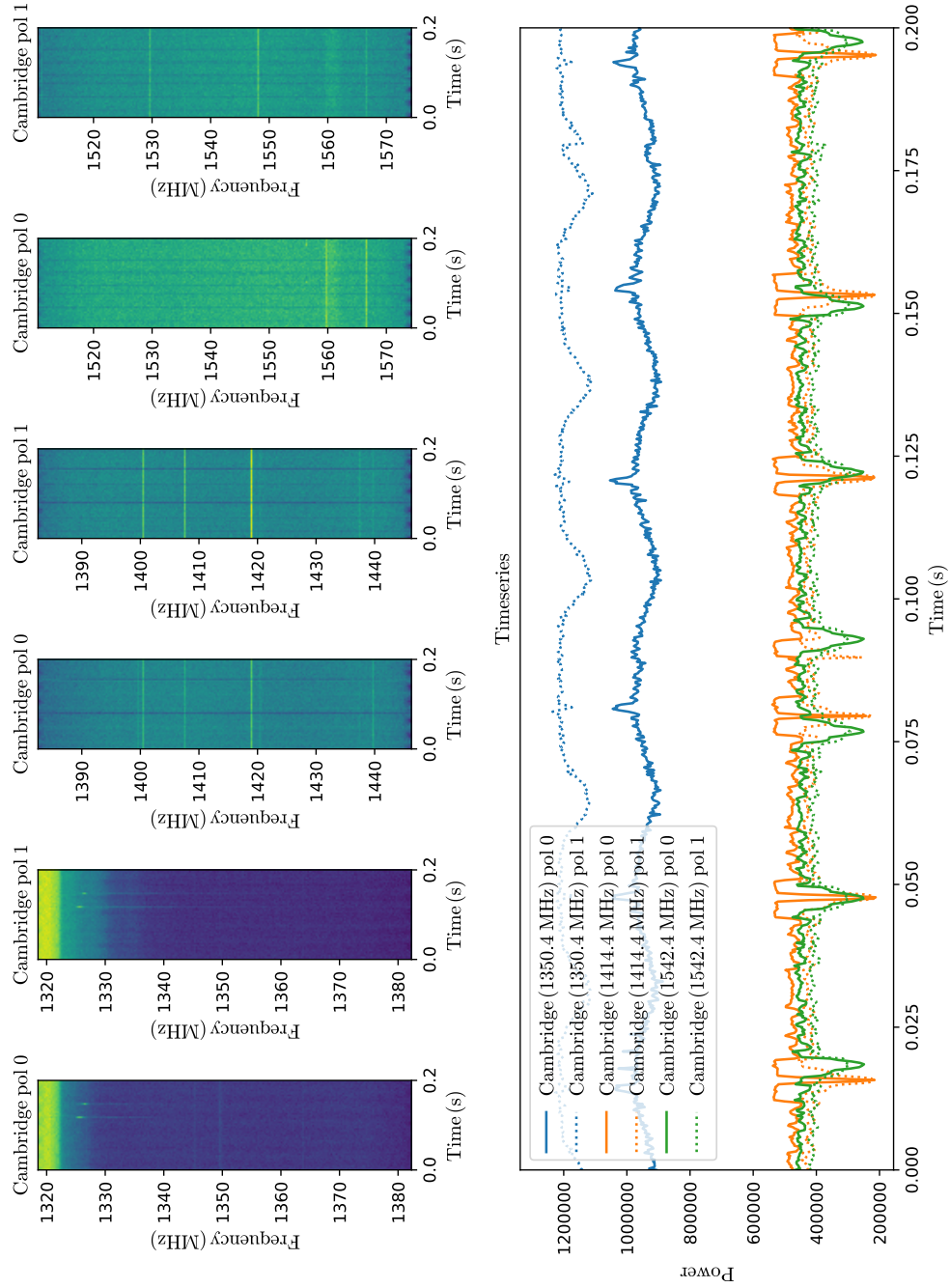
Figure 4.10: Plots in the style of Figure 4.7 for the individual polarisations of Cambridge's telescopes. Each dynamic spectrum is 200 ms in duration, with 250 ms temporal resolution and 250 frequency channels. Dynamic spectra are not bandpass subtracted.
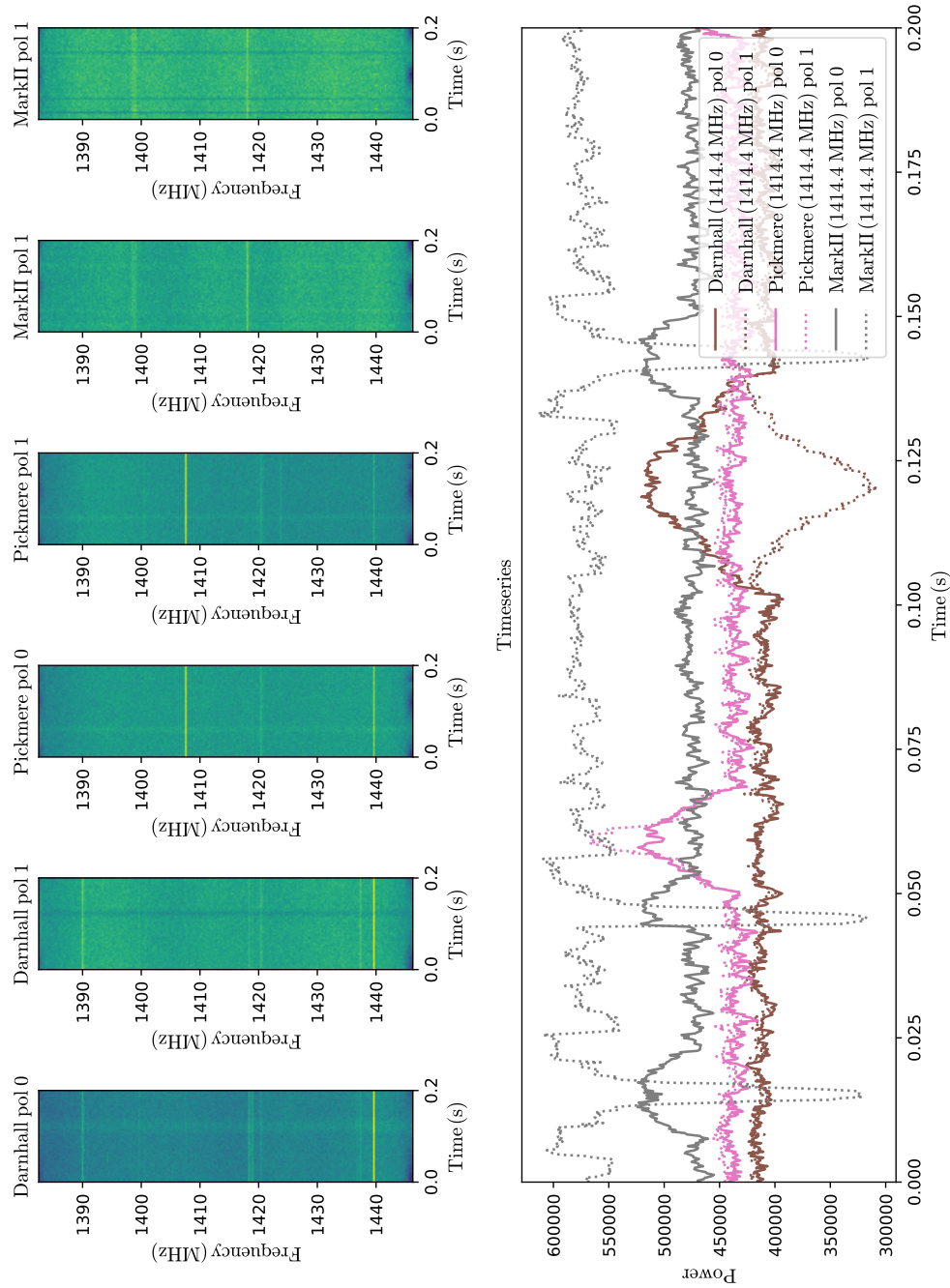
Figure 4.11: As Figure 4.10, comparing data from the Darnhall, Pickmere and Mark II telescopes.
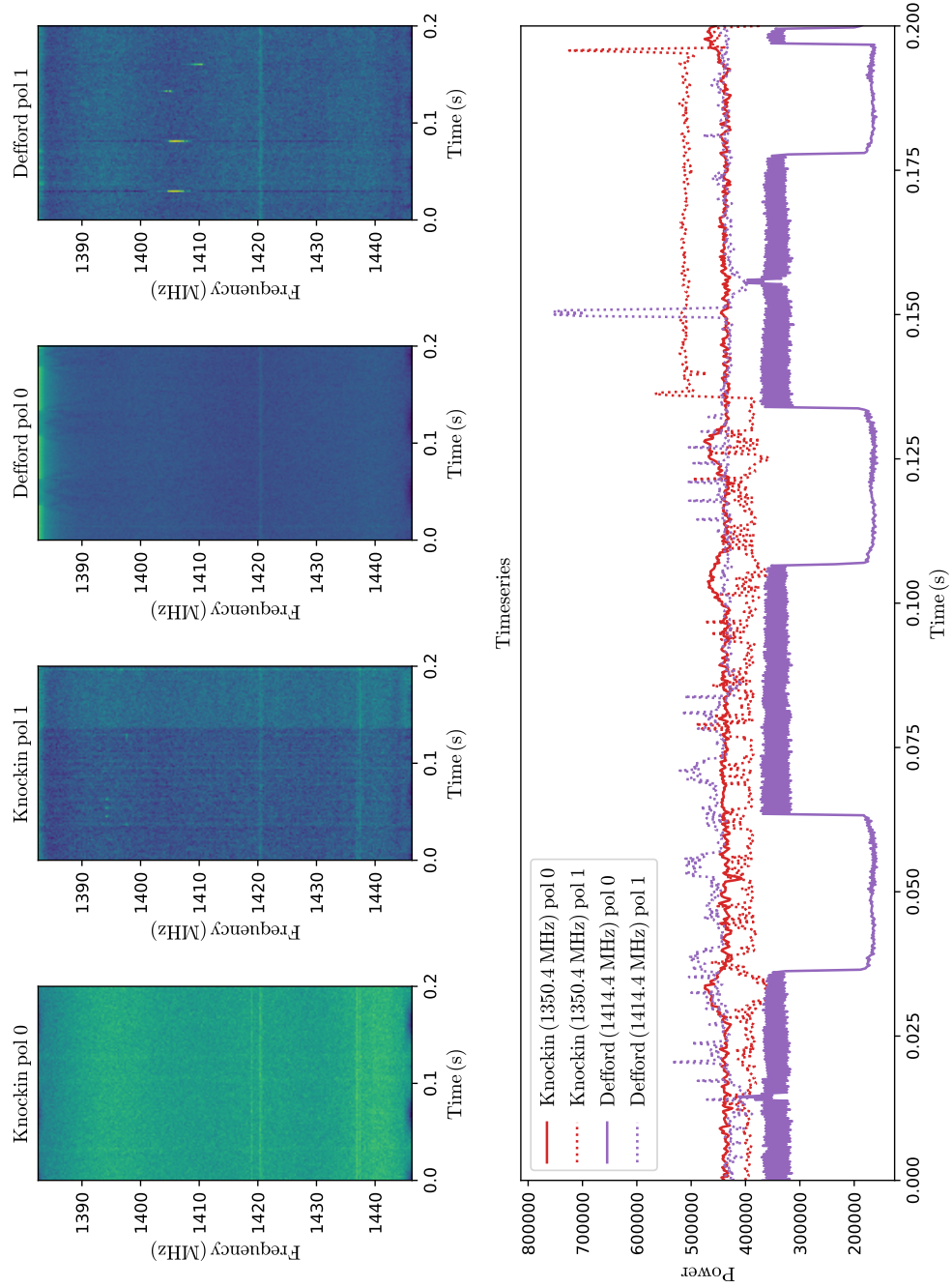
Figure 4.12: As Figure 4.10, comparing data from the Knockin and Defford telescopes.

Hardware	Quantity	Function	Path
Intel Core i7-5930K 3.5 GHz CPU	3 (1 per node)	offline data processing	N/A
Nvidea Geforce GT 730 GPU	3 (1 per node)	display	N/A
Nvidea Geforce GTX 980 Ti GPU	6 (2 per node)	real-time data processing	N/A
28 TB striped hard drive	3 (1 per node)	storage	`/stripe/`
512 GB SSD	3 (1 per node)	OS, home area and storage	`/state/partition1/`
48 TB NAS drive	1	archival storage	`/share/nas1/`

Table 4.4: List of hardware in the LOFT-e backend. Where appropriate (e.g. for data storage), system path is included.

Chapter 5

LOFT-e: Pipelines and Processing

In Chapter 4 the flow of data and metadata from e-MERLIN's telescopes to the LOFT-e backend was reviewed, and an example of e-MERLIN data which had been successfully captured and processed by LOFT-e was presented. The LOFT-e's data capture pipeline underpins e-MERLIN's high-time-resolution observing mode, and will be used both for commensal research and for targeted observations. Therefore in this chapter the data capture pipeline co-developed by the author is reviewed, and instructions for its implementation are outlined.

Several stages of processing must be applied to e-MERLIN's 64 MHz-wide data streams after capture, and before fast transient events may be optimally detected and studied. These stages form the LOFT-e data processing pipeline, developed by the author. This pipeline, as applied to data on a stream-by-stream basis, is also detailed here. Processing includes the use of several techniques for mitigating the sensitivity-reducing effects of RFI common to e-MERLIN, for example sporadic environmental events or persistent, periodic instrumentational interference (see Section 4.5). These techniques are discussed alongside examples of their application. LOFT-e's strategy for improving sensitivity via the combination of multiple frequency sub-bands and incoherent beamforming of dishes is then detailed. Variants on standard incoherent beamforming which may potentially be used to reject heavily RFI-contaminated data from individual dishes are introduced.

The techniques by which LOFT-e detects fast radio transient phenomena may be implemented at multiple points in the data processing pipeline: on individual dishes or frequency sub-bands, or on combined data products. These techniques are overviewed towards the conclusion of the chapter. Finally, areas of LOFT-e pipelines undergoing active improvement or in need of future development are reviewed.

Much of LOFT-e's software library has been purpose-built and extensively iterated upon over the course of the author's candidacy. The author clarifies where new software has been developed, or where existing software has been utilised where appropriate.

5.1   LOFT-e data capture

Initially data are streamed in 2-bit, 64 MHz-wide frequency sub-bands from the e-MERLIN correlator to LOFT-e via its baseline boards (see Section 4.4). A LOFT-e user may request a specific distribution of telescopes and sub-band centre frequencies to LOFT-e computing nodes according to their scientific aims, or accounting for expected RFI environments or telescope functionality at the time of the observation. Specific setups are decided by liaising with e-MERLIN staff and schedulers[1].

The resulting data and metadata are captured by the LOFT-e data capture pipeline. This software forms the basis of a new e-MERLIN high time resolution observing mode. At the time of writing, the LOFT-e data capture pipeline is specifically geared towards storing large quantities of data during targeted observations. Post processing (see Section 5.2) is then performed offline. In the future, when all necessary processing stages are efficient enough to operate in real time, the pipeline will be modified to enable processing of commensal observations[2]. Presently, this strategy enables the high time resolution analysis of known sources (e.g. pulsars and RRATs) with e-MERLIN. Here the author reviews the stages implemented by the data capture pipeline. A guide for its general use is provided in Appendix A.1. Results of such observations are presented in Chapter 6,

The data capture pipeline ensures all data and metadata necessary for fast transient analysis is being recieved by LOFT-e, then records output data products and observing logs in a structured format on LOFT-e's computing nodes. All software used in the data capture pipeline was developed by the author of this thesis alongside fellow PhD student, T. Scragg. A state machine description of the end-user data capture control program is shown in Figure 5.1[3]. An in-depth summary of the pipeline's process is provided in the following section.

---

[1] At the time of writing: schedulers.jb.man.ac.uk; Paul Harrison (paul.harrison@manchester.ac.uk).

[2] Via removal of the observation window check, reliance on source name and demanded azimuth and altitude coordinates, and output into filterbanking in the GPU memory in a ring buffer.

[3] At the time of writing, the current iteration of the end-user program is: `/home/cwalker/loft_e/lofte_receivers/eMControl_11cam.py`.
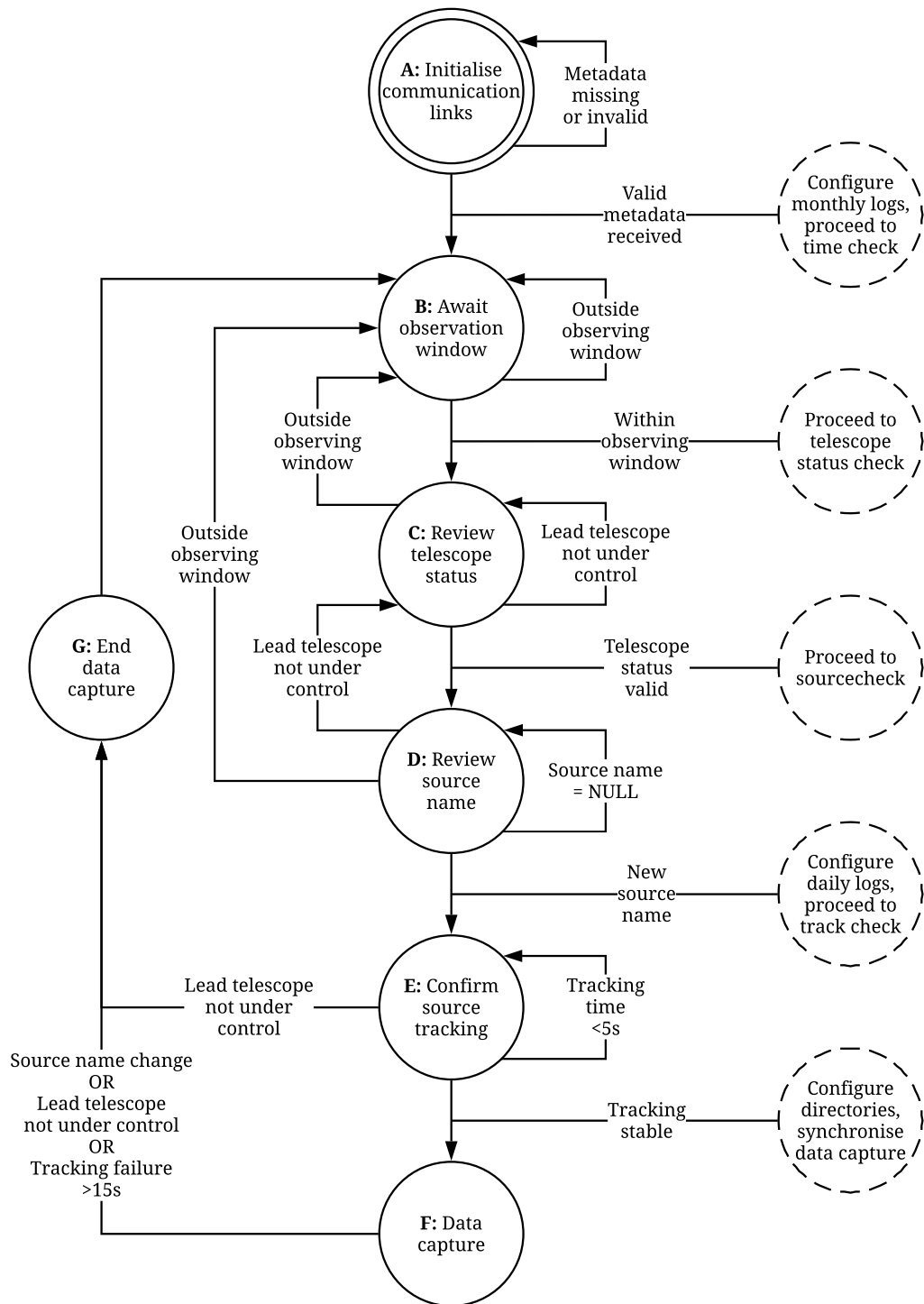
Figure 5.1: State machine description of the LOFT-e data capture pipeline. Data capture occurs during designated observation window after a series of telescope and VDIF export metadata checks. Diagram based on figure by T. Scragg.

5.1.1   The data capture pipeline

In this section the stages undertaken by the data capture pipeline, represented pictorially by Figure 5.1, are detailed. Upon activation of the pipeline, monthly system log files are created.[4] These log files record terminal output from the data capture pipeline. If the pipeline is run multiple times in one day, output will be appended to the existing log file. Following creation of log files, communication links between LOFT-e and e-MERLIN are initialised and a series of checks are performed. Details of the checks are as follows:

Initially, a first pass loop ensures valid telescope and VDIF export metadata are being recieved (see Figure 5.1: A). Copies of this metadata are stored to the system logs[5]. When all necessary metadata are being received, the main control routine begins.

Within the main control routine the current Coordinated Universal Time (UTC) according to LOFT-e's head node is compared to an observing window of user-designated start and end times (see Figure 5.1: B). When the current time falls within the specified observing period, the status of the user-designated lead telescope is obtained from telescope metadata. If the lead telescope's azimuth and altitude motors are running and the instrument is under JBO automatic control (see Figure 5.1: C), the demanded observation source name is obtained (see Figure 5.1: D). Individual source log files are then created.[6] Finally, a comparison between the lead telescope's current and demanded azimuth and altitude co-ordinates ensures the lead telescope is on source (see Figure 5.1: E).

When the lead telescope is under JBO control with active motors, and current co-oordinates lie within a certain accuracy of those demanded the entire e-MERLIN array is assumed to be functional, on-source and tracking correctly. Data capture is then initialised (see Figure 5.1: F)[7].

Data directory paths are synchronised across all LOFT-e computing nodes according to the date, UTC, and source of observation[8]. A new data directory is created each

---

[4]Monthly log files are stored on individual computing nodes under the naming convention: `/stripe/cw/loftelog/YYYYMMDD/LofteSysLogYYYYMMDD.log` where YYYY, MM, DD correspond to the year, month and day of observations respectively.

[5]Observing schedule metadata is currently compulsory for filterbanking but not compulsory for data capture. Its output is not currently stored to logs, however this should be amended in case of loss.

[6]Source log files are created according to the naming convention:
`/stripe/cw/loftelog/YYYYMMDD/YYYYMMDD[source name]/YYYYMMDD[source name].log`.

[7]At the time of writing, directory initialisation and data capture is performed by program: `/home/cwalker/loft-e/lofte_receivers/run_lofte_temp_7t.py`.

[8]An example of data directory naming conventions is:
`/stripe/cw/2018-06-26_18:11:09.049494-UTC-407+284/`.

time e-MERLIN changes source, or if a sufficient period of tracking failure occurs on the lead telescope. Within individual data directories, VDIF export and observing schedule metadata are recorded according to the naming conventions established in Chapter 4 for use in LOFT-e's data processing pipelines.[9]

Initially, a single packet of data from one stream is captured on the LOFT-e head node, `transient`. Data recording is synchronised by reading the timestamp for this packet and beginning capture on all streams 30 s later. A single baseline board provides two interleaved 64 MHz data streams for a single telescope and frequency sub-band, one per polarisation. Polarisations are recorded individually. Data files themselves are named according to the Modified Julian Date (hereafter MJD) at the time of recording, the ID of the baseline board of origin, and polarisation. Within a data directory, a single data file may store up to 240,000 VDIF frames, corresponding to 10 minutes of data. For observations longer than 10 minutes, multiple files are created and distinguished with a file extension number[10].

Telescope metadata for the user-designated lead telescope is routinely monitored and appended to system logs while the pipeline is active. If fast radio transient candidates are detected during post-processing, this enables cross-referencing between their timestamps and the on-sky co-ordinates of e-MERLIN at the time of observation. If during observations a check is failed, data capture ceases and the pipeline returns to an earlier stage in the checking process (see Figure 5.1: G).

5.1.1.1   Accounting for data loss

Data is transmitted to LOFT-e via User Datagram Protocol (UDP) packets. During data capture, packets may be lost due to inadequate hardware write speeds or overflow of intermediary buffers between e-MERLIN and LOFT-e's data receiving sockets. Examples of significant packet loss encountered during early data capture tests are shown in Figures 5.2 and 5.3.

In the general six-telescope, twelve baseline board LOFT-e configuration, data rates reach 256 MB/s per computing node. Figure 5.2 clearly illustrates the effect of insufficient write speeds. During this experiment, four input data streams (for a total data rate of 256 MB/s) were directed to a single computing node and written to a 4 TB

---

[9]At the time of writing, observing schedule metadata are only stored to the head node. As they must therefore be copied to the other nodes, human error may result in correct schedules being accidentally overwritten. This is easily fixable by changing `/home/cwalker/loft-e/lofte_receivers/run_lofte_temp_7t.py`.

[10]An example of the data file naming convention is: `mjd_58296.3633083_data_13000.pol0.File1`.
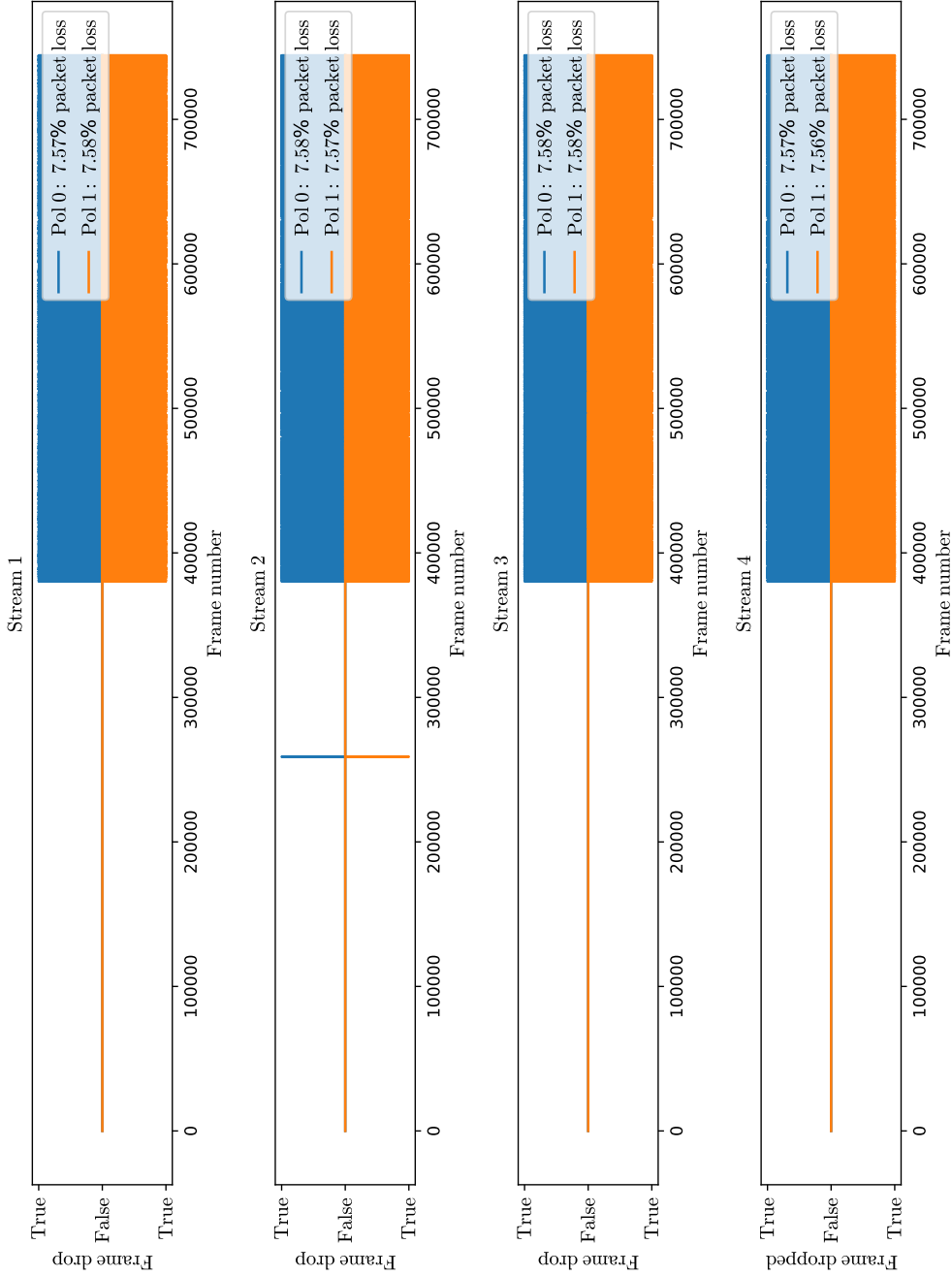
Figure 5.2: Results of data drop tests for four concurrent, three-minute data streams to a single LOFT-e computing node. The e-MERLIN correlator broadcasts a data stream as 'frames' at a rate of 4000 frames per polarisation second. Each 8032-byte frame contains 4000 time samples. Therefore total data rate here is 256 MB/s. Data is being written to a 4-TB striped hard drive with a write speed of 238 MB/s. The inadequate write speed results in ~ 8% loss of data over the observing period. Dropped frames are indicated by a value of 'True' on the y-axis.
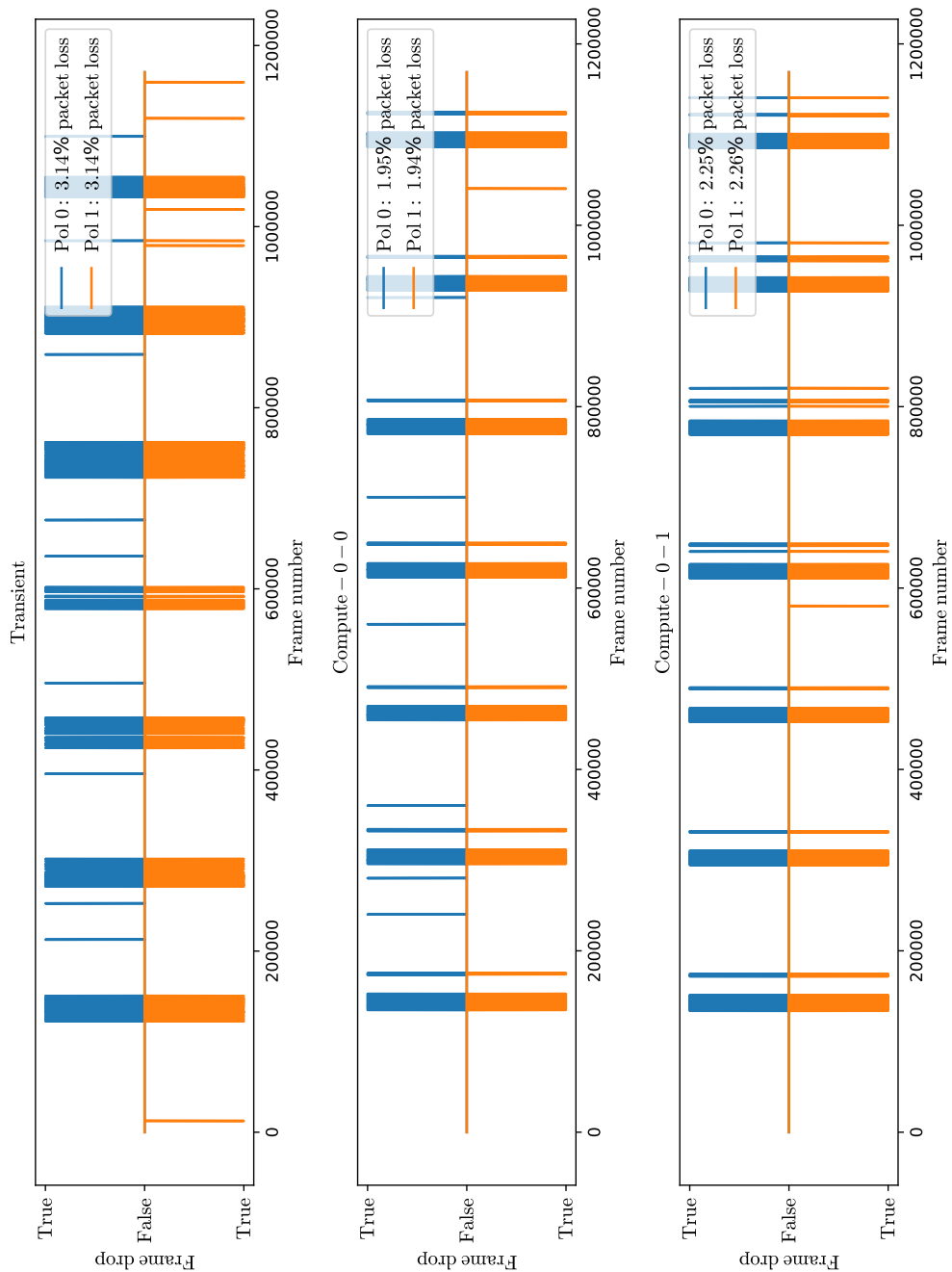
Figure 5.3: As Figure 5.2, for data drop tests for a single, three-minute (120,000 frame) data stream per LOFT-e computing node being written to a shared 48-TB SSD drive. The total data rate per node is 192 MB/s. Despite the SSDs' adequate write speed (407 MB/s) non-negligible data loss occurs.
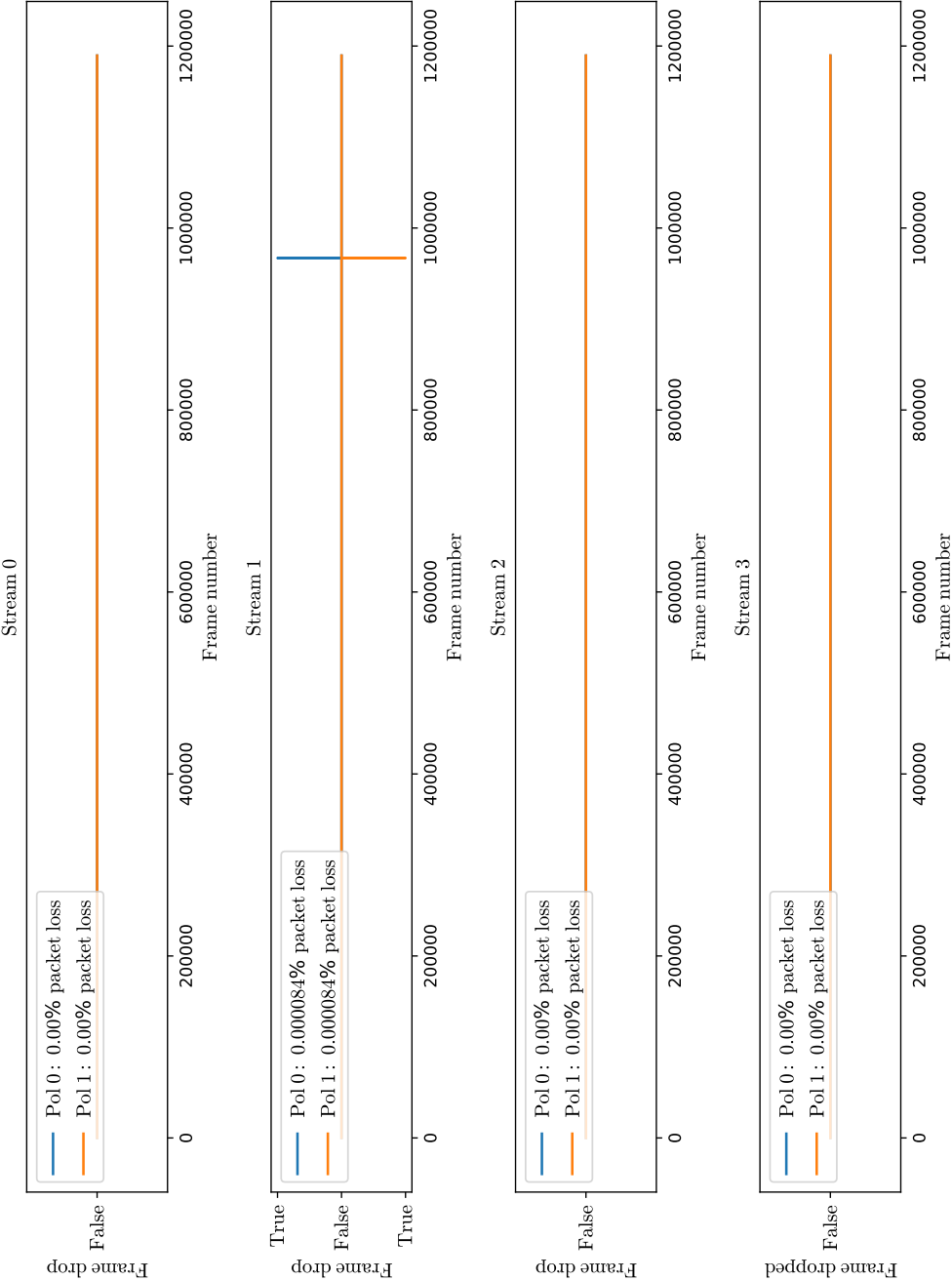
Figure 5.4: As Figure 5.2, for data drop tests for four concurrent, three-minute (120,000 frame) data streams to a single LOFT-e computing node. The total data rate is 256 MB/s. Data is being written to a 28-TB striped hard drive with a write speed of 383 MB/s, resulting in minimal packet loss. This setup is standard LOFT-e practice for a six-telescope, twelve baseline board e-MERLIN configuration during non-commensal data storage. Occasional packet loss still occurs (see stream 2) which is filled in real time by the LOFT-e data capture pipeline.

striped hard drive with a write speeds of 238 MB/s. This hardware configuration resulted in up to 8% packet loss. Figure 5.3 shows results of a second experiment, where a single input data stream per computing node (for a total data rate of 196 MB/s) was written to LOFT-e's shared 48 TB solid state disk (SSD) drive with a write speed of 407 MB/s. Despite adequate write speeds, packet loss during this experiment was also non-negligible ($\sim 3\%$) due to inefficient data packing on the SSD during the writing process.

To reduce potential signal loss during observations, in dropped data should be minimised. Despite SSD drives being less prone to failure with age, 3% packet loss was deemed unacceptable for real-time data storage. Therefore LOFT-e data capture protocol was modified to data streams to 28 TB striped hard drives with 383 MB/s write speeds, resulting in significantly less packet loss, as shown in Figure 5.4. This hardware configuration is more prone to failure, and any stored data must be archived regularly.

As evidenced by Figure 5.4, despite adequate write-speeds, buffer issues may still result in small numbers of packet drops. This effect varies on a stream-by-stream basis, potentially resulting in previously synchronised data from two telescopes losing synchronisation. Unsynchronised data will result in sub-optimal sensitivity improvements during beamforming. Additionally, any lost packets will leave an otherwise evenly sampled data time series non-uniformly spaced. This violates the requirement of uniformly spaced input data for FFTs (which are used during filterbanking, see Section 4.2.2). It is therefore necessary to account for missing packets during the data capture process.

The data capture pipeline has therefore been designed to identify and fills in missing packets in real time. To do so, the time of each incoming VDIF frame is calculated[11], and missing frames are filled with a copy of the preceding frame. Each filler frame is updated with a timestamp-corrected header to avoid potential issues during beamforming[12].

---

[11]Times are calculated using two parts of the VDIF header: a frame number, running from 0-3999, and the number of seconds past since a reference epoch (SFRE), beginning at UT 00:00 on 01/01/2000.

[12]At the time of writing, data are written by:
`/home/cwalker/loft-e/lofte_receivers/receive_data_8.c`.

5.2   LOFT-e data processing

As discussed in Chapter 3, optimising the incoherent sensitivity of an interferometer for fast transient detection involves maximising the number of telescopes and total bandwidth used in observations, along with optimisation of the sampling interval (according to the width of individual pulses) and observation time (when averaging multiple periodic pulses during pulsar observations). Therefore LOFT-e applies several stages of processing to captured data.

Initially, individual data streams are filterbanked, providing finer spectral resolution and increasing the data's sampling interval. Various RFI mitigation techniques are then applied to filterbanks on a stream by stream basis. All output filterbanked data products then undergo beamforming on the LOFT-e head node on a sub-band by sub-band basis. Ideally, all telescopes involved with the observation will be used, however variants on standard incoherent beamforming which allow rejection of data from RFI contaminated dishes may be utilised. Finally, all sub-bands for an observation are combined, maximising the bandwidth of the data to be searched for fast transient phenomena.

The data processing pipeline[13] has been modularly designed and is currently semi-automated, allowing for analysis of data from individual telescopes and frequency sub-bands at each stage of the process during LOFT-e's commissioning phase. Fast transient detection (see Section 5.3) may be optionally performed at several intervals. This has proved useful for analysing the effects of each stage on e-MERLIN's sensitivity during pulsar observation tests. These tests will be reviewed in Chapter 6. LOFT-e's current data processing pipeline is illustrated by the flow diagram in Figure 5.5. The relevant stages are detailed in this section. The author explicitly notes where user input is required.

5.2.1   Filterbanking

During filterbanking of an evenly spaced voltage data time series (as is provided by e-MERLIN), N time samples ($t_0, t_1, ...t_{N-1}$) are accumulated and an N-point discrete Fourier transform (DFT):

$$F_k = \sum_{n=0}^{N-1} t_n \exp\left(-2\pi i n k/N\right), \tag{5.1}$$

---

[13]The data processing pipeline is found in: `/share/nas1/LOFT-e/software/`.
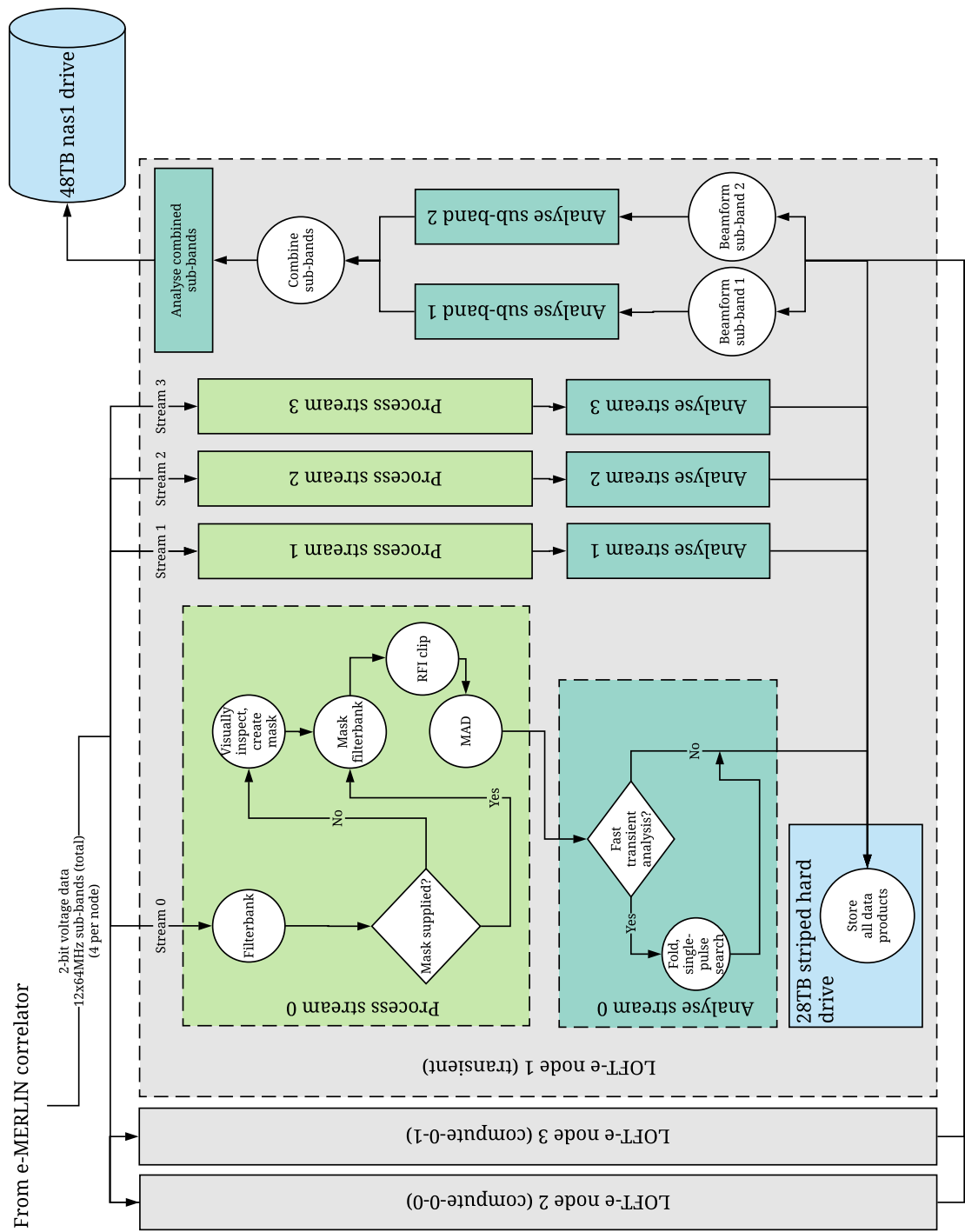
Figure 5.5: The LOFT-e data processing pipeline for general six-telescope, twelve baseline board e-MERLIN configuration. Four data streams are processed separately on each computing node. Data products are then transferred to the LOFT-e head node, transient, for combination and further analysis. Final data products are stored on a 48-TB nas1 drive.

is performed on the data, breaking them down into N complex exponential Fourier components ($F_0, F_1, ...F_{N-1}$). In this equation $i = \sqrt{-1}$ and each Fourier component, F_k, corresponds to the signal at frequency ν_k:

$$\nu_k = k/(Nt_{samp}), \tag{5.2}$$

where t_{samp} is the original sampling time (Lorimer & Kramer, 2012). After filterbanking, the new sampling time becomes $(N \times t_{samp})$ and the bandwidth of each frequency channel becomes $1/(N \times t_{samp})$. During filterbanking, the $k = 0$ Fourier component $F_0 = \sum_{n=0}^{N-1} t_n$ corresponds to the average power in the band. This is known as the DC component, and is discarded.

Filterbanking within the LOFT-e data processing pipeline is achieved using fast Fourier transform (FFT) algorithms which perform DFTs efficiently[14] The output data products are stored in the commonly used two-dimensional sigproc filterbank format.[15] known as a dynamic spectrum. A dynamic spectrum stores the Fourier components of n_{chan} frequency channels ($F_1, F_2, ...F_{n_{chan}}$) for n_t times ($t_1, t_2, ...t_{n_t}$), and may be mathematically described as $D_s(t_j, F_k)$. From a dynamic spectrum, a one-dimensional time series may be produced by summing over all frequency channels:

$$T_i = \sum_{j=1}^{n_{chan}} D_s(t_i, F_j). \tag{5.3}$$

Such time series are used to search for fast radio transients (see Section 5.3). By summing over all times, an average bandpass may also be created:

$$B_j = \sum_{i=1}^{n_t} D_s(t_i, F_j). \tag{5.4}$$

By accumulating more time samples, data products with narrower frequency channels are created, allowing the fine spectral behaviour of transients to be studied by allowing for more accurate removal of the effects of dispersion (see Chapter 2) via incoherent dedispersion (see Section 5.3.1). Filterbanking to fine spectral resolutions also allows the spectral behaviour of transients to be studied in detail. Additionally, to reduce memory requirements, filterbank files are often averaged in time, frequency, or

---

[14]In $N \log_2 N$ operations (Lorimer & Kramer, 2012).
[15]`sigproc.sourceforge.net` (Accessed: 25/09/2018)
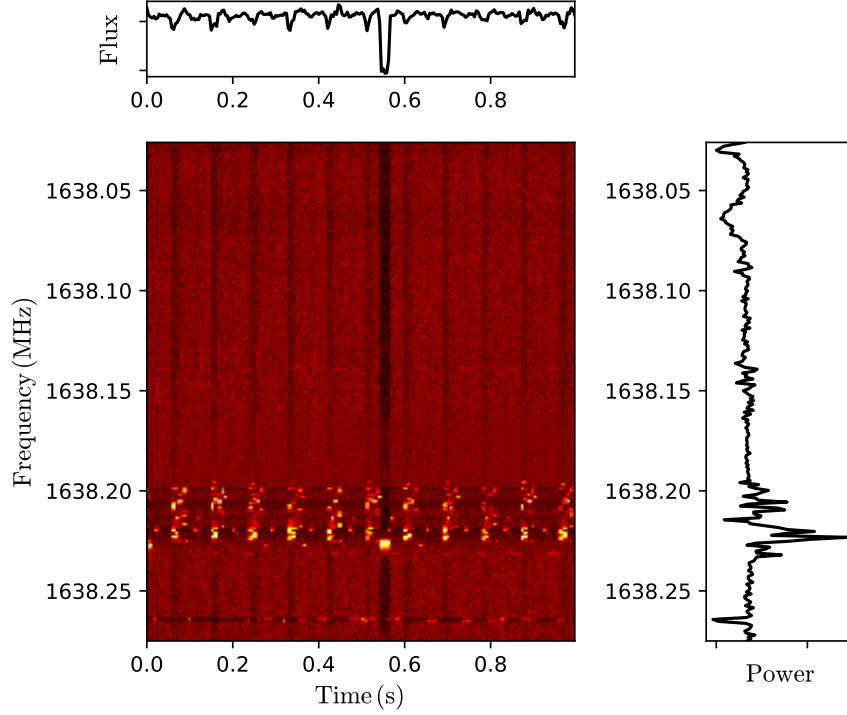
Figure 5.6: Dynamic spectrum (center) in filterbank format, with time series (top) and band-pass (right). Data was taken during an observation of PSR B1935+1616 at L-band and integrated to 4.16 ms. Repeat impulsive RFI and persistent broad-band RFI are visible.

both. An example LOFT-e sigpyproc filterbank, complete with time series and band-pass, is shown in Figure 5.6.[16]

Sensitivity to transients also depends on the data's sampling interval, so the trade-off between spectral and temporal resolution must be considered. Sampling intervals should not be wide enough to allow significant dispersive smearing within channels. The ideal sampling interval is equal to the intrinsic width of a burst (see Section 3.1.2). As fast transient widths range from sub-millisecond to tens of milliseconds (see Section 1.3.1), LOFT-e filterbanks are currently hard-coded to 64 μs temporal and 256×0.25 MHz spectral resolutions. If further time/frequency averaging is required, this may be done externally[17].

---

[16]Bandpasses should ideally be flat.

[17]Using, e.g. sigpyproc: `https://github.com/ewanbarr/sigpyproc` (Accessed: 25/09/2018).

LOFT-e data products are produced in sigpyproc filterbank format[18]. This format requires the generation of headers which include information about the observations[19]. Bandwidth and observing frequency information are extracted from VDIF export metadata. Start time information is extracted from the header of the first relevant data file stored in the directory[20]. Target source and sky co-ordinate details are extracted by cross-referencing times with observing schedule metadata.

Filterbanking may be performed in parallel across LOFT-e's computing nodes, however on a single node, within a single directory, filterbanking is performed on a stream-by-stream basis.[21] Therefore filterbanking of individual telescopes, sub-bands, and data files is not performed in parallel. Previous Python-based filterbanking software has been superceded by GPU-based software which processes data faster than real-time[22]. During filterbanking[23], both polarisations are combined into a Stokes I (total intensity) data product, and output files are stored in the same folders as their raw voltage data products.

Currently, filterbanking is run by an end-user script[24] which calls the main LOFT-e data processing pipeline[25] Within this script, file extension numbers to be filterbanked (if observations were > 10 min in duration) are specified. As the observing schedule metadata received by the pipeline may contain multiple schedules, the correct observing schedule index must also be provided. Filterbanks may be output in either 32-bit unsigned (to avoid loss of precision due to scaling) or scaled 8-bit unsigned (to reduce output file size) format. The desired output format must therefore be specified[26].
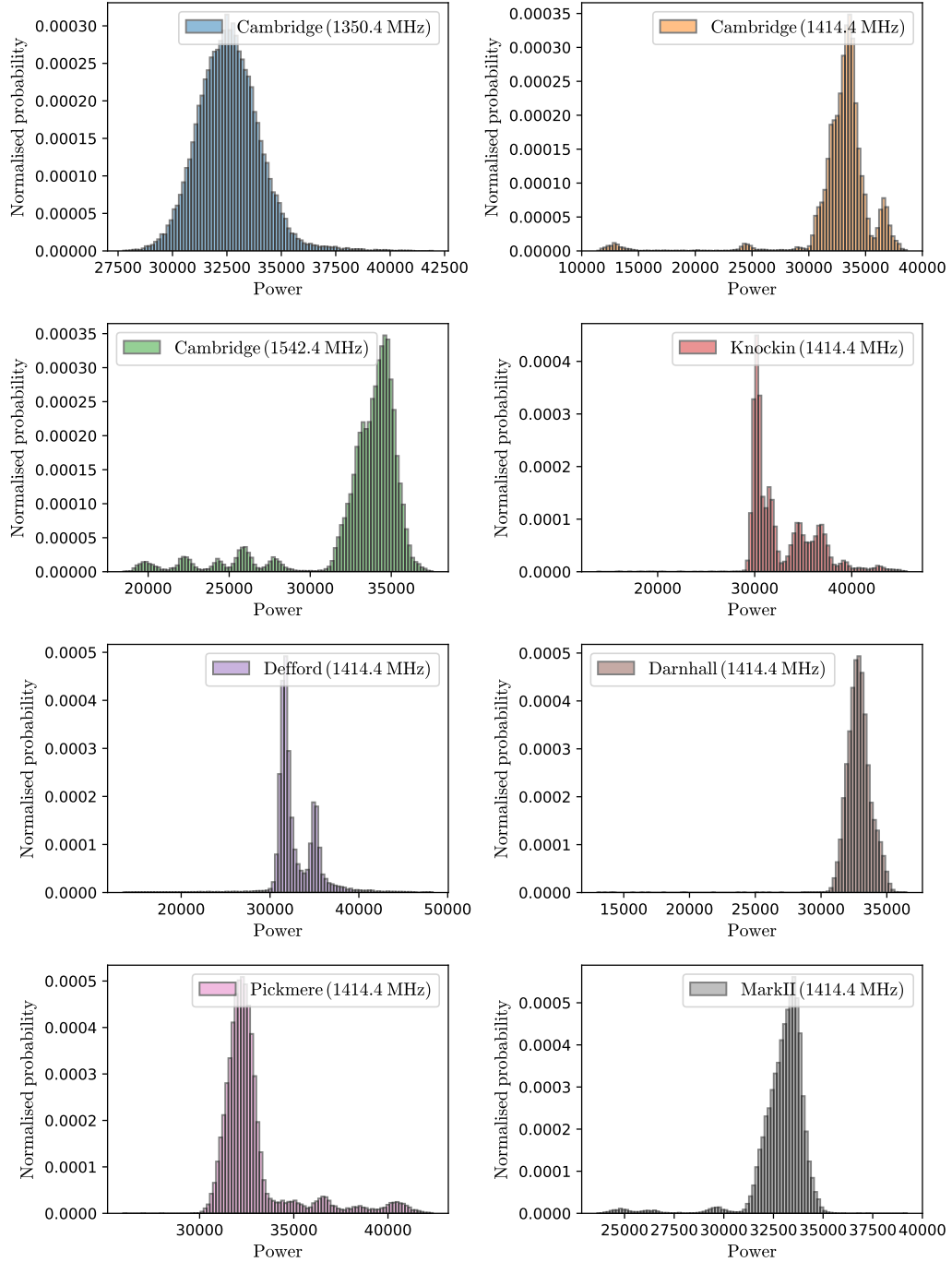
Figure 5.7: Histograms of counts against time series power for one-second-duration time series data acquired from all e-MERLIN elements during the observation setup used for Figures 4.7 and 4.8. Histograms have been normalised by total power. Data from each telescope shows a deviation from a Gaussian distribution which has been induced by RFI.

5.2.2   Data stream-based RFI mitigation

As discussed in Chapter 2, the ensemble emission of astrophysical radio sources results in a background of Gaussian random noise. Radiometer theory relies on this assumption, as does digitisation theory (Dicke, 1946; Burke & Graham-Smith, 2002). Signal to noise thresholds used to identify fast transient events in data ideally also make use of this assumption (Cordes & McLaughlin, 2003).

RFI is ubiquitous at radio observatories, and is not Gaussian in nature. Extreme forms of RFI include persistently contaminated frequency channels (see, e.g., Figure 5.6), or short impulsive bursts which cover broad broad bandwidths. Impulsive RFI may be periodic or quasi-periodic, mimicking the behaviour of naturally repeating signals such as pulsars. Examples of particular relevance to e-MERLIN are an 8 s periodic airport radar often detected by the Lovell and Mark II telescopes (see Section 6.1.1), and the $\sim$ 25 Hz signal identified in Section 4.5.

Fundamentally, the presence of RFI causes significant departures from Gaussianity in data. Figure 5.7 presents normalised 100-bin histograms of one second of time series data for each e-MERLIN telescope. These data are from the same set of LOFT-e observations discussed in 4.5 and illustrated in Figures 4.7 and 4.8. Each profile displays a deviation from the Gaussian profile expected for RFI-free data. Knockin (1414 MHz), Defford (1414 MHz), Pickmere (1414 MHz), Mark II (1414 MHz) and Cambridge (1414 MHz, 1542 MHz) sub-bands display multi-peaked distributions. Darnhall (1414 MHz), Mark II (1414 MHz) and Cambridge (1414 MHz, 1542 MHz) sub-bands display leading tails, with many time samples having lower powers than the average value. These distributions arise due to RFI which periodically dips below the time series average power. Conversely Knockin (1414 MHz), Pickmere (1414 MHz) and Cambridge (1414 MHz) sub-bands multiple trailing peaks or heavy tails, owing to periodic RFI

---

[18]Lorimer (2011): `http://sigproc.sourceforge.net`(Accessed: 25/09/2018).

[19]Headers are created using:
`/share/nas1/LOFT-e/software/analysis_pipeline/make_gpufil_headers_3.py`,
`/share/nas1/LOFT-e/software/analysis_pipeline/jsontoheader_2.py`.

[20]Therefore it is noted here that data capture synchronisation across computing nodes as discussed in the previous section is critical for further data processing.

[21]Filterbanking is performed by:
`/share/nas1/LOFT-e/software/analysis_pipeline/fast_fb_lofte_tom3.py`

[22] `/home/malenta/code/gpufilterbank/vdif_new`

[23]`/share/nas1/LOFT-e/software/analysis_pipeline/fast_fb_lofte_tom3.py`

[24]E.g. `/share/nas1/LOFT-e/software/makefilscript_100818`.

[25]Most recent version: `/share/nas1/LOFT-e/software/analysis_pipeline/LA_pipeline_14.py`.

[26]During 8-bit scaling process, the mean and standard deviation used are obtained from the first 5 s of data. Any values falling outside of the unsigned 8-bit range (i.e. <0 or >255) are clipped.

which rises above the average time series power.

RFI may present a variety of problems during fast transient detection. Impulsive RFI bursts may mimic signals from sporadically emitting sources (e.g. RRATs) which may compromise our ability to measure their period derivatives and derive phase-connected timing solutions. Periodic RFI may be flagged by periodicity searches, and bright, temporally narrow broad-band single pulses may be detected during single-pulse searches. Such scenarios may significantly increase the number of fast radio transient candidates to be examined by eye. At its most extreme, RFI may call into question the astrophysical nature of real transient phenomena. Perytons, first recorded by Burke-Spolaor et al. (2011) are an example of impulsive RFI with an apparent DM sweep. Peryton's apparent similarity to FRBs called into question the latter's authenticity until their nature as emission from prematurely opened microwave ovens was established (Petroff et al., 2015b).

Digitisation of extremely bright RFI may saturate entire frequency channels or time samples in data products, rendering them unusable. The presence of RFI in certain e-MERLIN frequency sub-bands severely impacts recovered S/Ns during pulsar observations (see Chapter 6). It is therefore important to remove as much RFI as possible before fast transient detection is attempted. Post-filterbanking, the LOFT-e data processing pipeline utilises a variety of RFI mitigation techniques, from brute force masking of persistently bad channels to more complex methods taking advantage of multiple available telescopes. Techniques applied to individual data streams before combination are described in detail in this section.

5.2.2.1 Masking

The presence of persistent RFI may render the band unusable. These contaminated channels may reduce S/Ns of astrophysical signals in radio data. While this may occur on a timescale similar to the length of a single observation, it may also occur on much longer timescales, effectively becoming permanent features of e-MERLIN's observation environment. An example of persistent RFI results from the periodic signal identified in Section 4.5.

Figure 5.8 shows the upper portions of filterbanked data products taken at different times during an observation for two Cambridge frequency sub-bands and one Knockin frequency sub-band. Three different sets of times are shown for each telescope and sub-band. For each telescope/sub-band combination, a 'striping' behaviour is seen across the upper ~ 8 frequency channels of the filterbank file. This feature is consistent
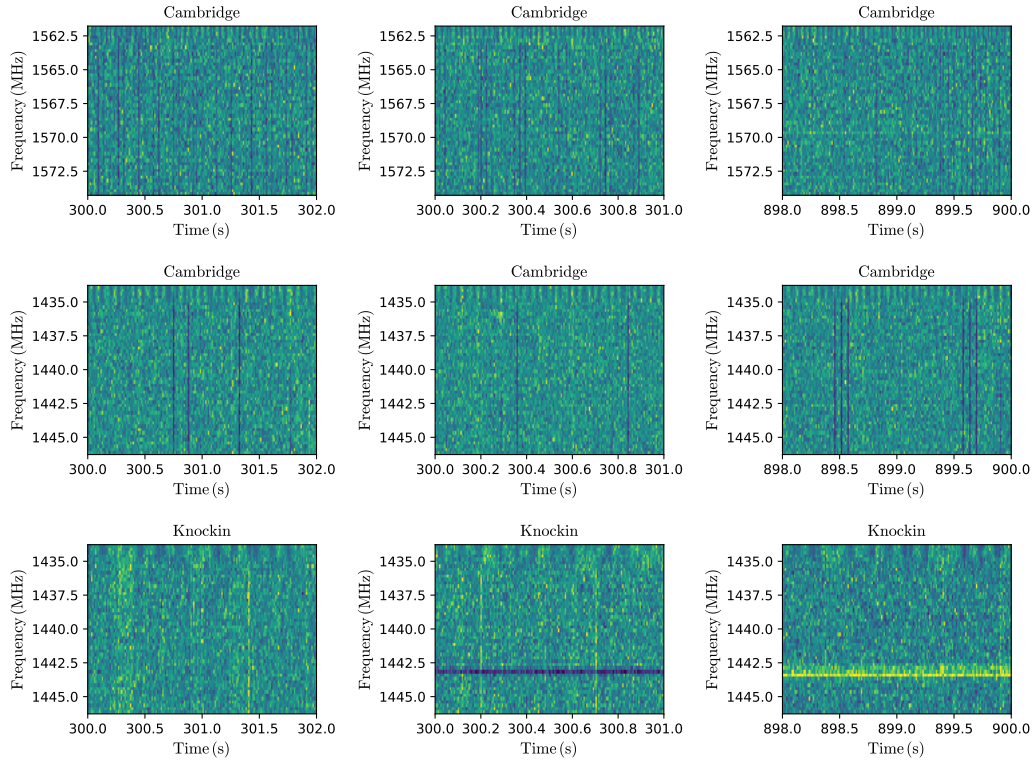
Figure 5.8: Dynamic spectra for upper portions of e-MERLIN sub-bands centered on 1542 MHz and 1414 MHz for Cambridge and Knockin telescopes. Dynamic spectra have 0.25 MHz spectral resolution and 64 μs temporal resolution. Data in each column is simultaneously recorded. Each row contains multiple sub-plots for the same telescope and frequency sub-band at different times. For each sub-plot, observing frequency is plotted against time since observations began. Colour indicates relative intensity.

throughout observations, as opposed to, e.g., the horizontal narrow-band RFI which appears in Knockin's latter two data sets. The exact period and width of this feature vary between telescopes and sub-bands, however the behaviour is visible in every e-MERLIN dish. The behaviour is additionally seen persistently at the bottom of each sub-band.

To remove persistent RFI features, masking may be employed. During masking, flagged filterbank frequency channels are completely excised of data and replaced with, e.g., an expected mean value for the data. The technique is usually employed before searches for fast transients take place to avoid accumulating large numbers of false candidates (Lorimer & Kramer, 2012). Mask files may be provided to LOFT-e's processing pipeline for each data stream captured during observations. Mask files consist of a list of numbers corresponding to frequency channels to mask prior to further analysis. Masking is performed[27] on copies of each filterbank file. All masked channels are replaced with the best-case-scenario mean value assuming Gaussian data and optimal scaling (i.e. 128 for unsigned 8-bit filterbanks). Output masked data products have amended `.masked.fil` file extensions.

Conversely, if the RFI environment is not well known, data processing may be optionally terminated after filterbanking. Commonly contaminated channels may then be found via visual inspection of filterbanked data products, masks may be created, and the pipeline may be restarted utilising the provided mask files. Mask files must be stored in a single directory and named for their corresponding baseline boards[28]. Tools for efficient mask creation have been developed for this purpose[29] and instructions for their use are provided in Appendix A.1.2.

5.2.2.2 Time-domain RFI clipping

Impulsive RFI may occur sporadically or periodically during observations. Figure 5.9 presents 0.2 second of data captured from Cambridge's sub-band centered on 1414.4 MHz. To visualise the broad-band nature of impulsive RFI within this observation, the 256-channel, 64 MHz-wide filterbank file has been averaged every 32 frequency channels, resulting in eight time series plots with different centre frequencies. A deviation from average power in the form of a dip is visible in each sub-band.

---

[27]Using `/share/nas1/LOFT-e/software/analysis_pipeline/mask_filterbanks_1.py` and `/home/cwalker/loft-e/lofte_receivers/RFIclean/pulsarTools/RFIclean`.

[28]E.g. `13000.mask`.

[29]`/share/nas1/LOFT-e/software/analysis_pipeline/psh_to_rfimask_lofte.py`
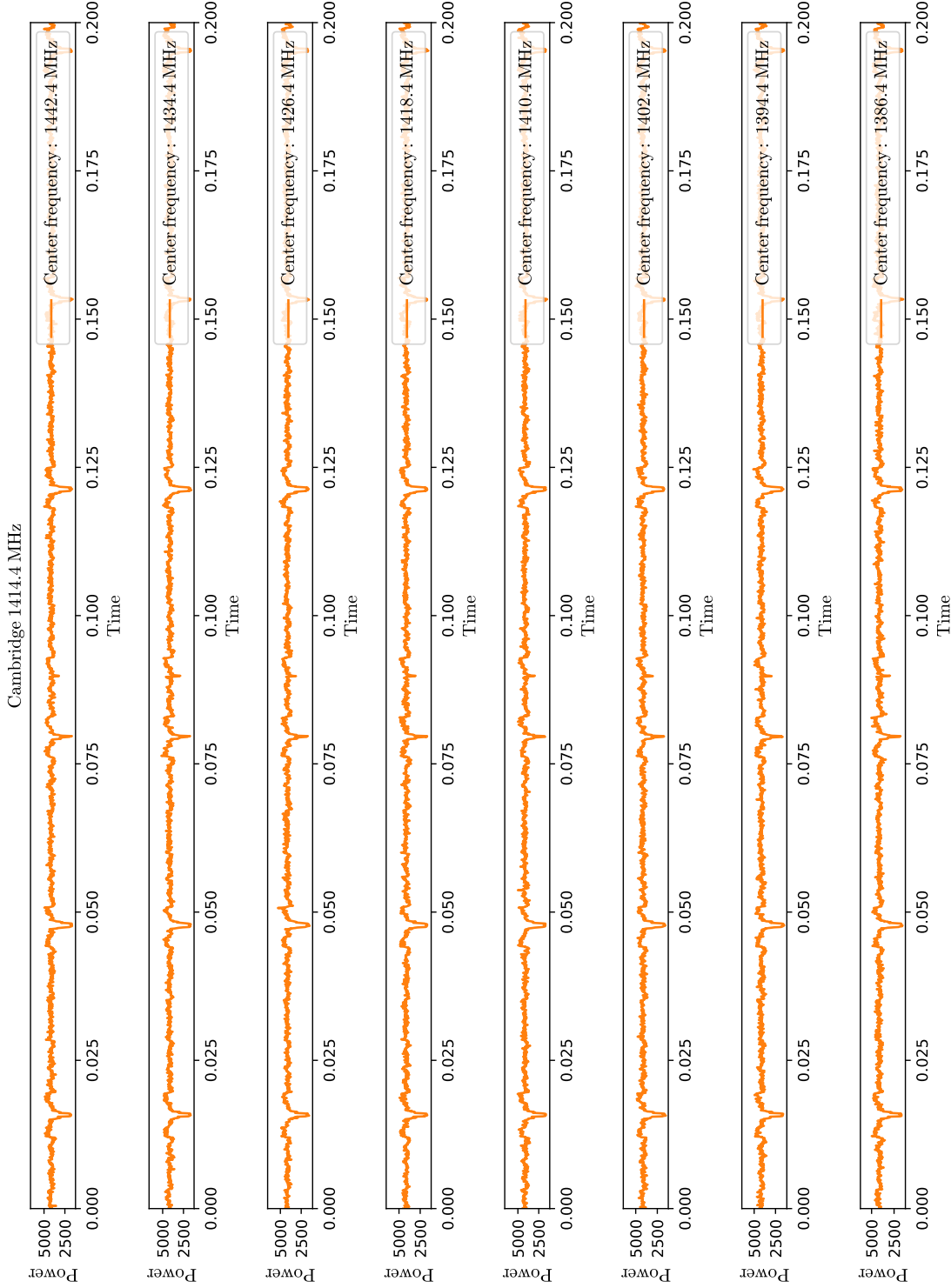
Figure 5.9: Frequency averaged power against time datasets acquired from the Cambridge telescope's 1414 MHz-centred sub-band during LOFT-e observations. A 0.2 s filterbank file has been averaged in frequency every 32 channels, resulting in eight time series plots at different center frequencies. Impulsive RFI in the form of periodic dips below the average power is visible throughout the sub-band.

The dips occurs periodically and contemporaneously across the entire bandwidth of the sub-band.

In order to mitigate the effects of broad-band impulsive RFI, a time-domain RFI clipping algorithm has been developed. This algorithm may optionally be used during the next stage of LOFT-e's processing pipeline. The stages of the algorithm are as follows:

1. **Read spectra**: Due to memory constraints clipping is performed on the entire dynamic spectrum (i.e. a filterbank format data file) by sequentially processing chunks of individual spectra (i.e. all frequency channels for one time sample). Currently, 40,000 spectra are read into memory at one time.

2. **Rescale frequency channels**: For a set of spectra, the median and standard deviation of each frequency channel are calculated. These values are obtained using an iterative median clipping algorithm[30]. Frequency channels are then rescaled by subtracting the median and dividing by the standard deviation of the channel.

3. **Perform time-domain thresholding**: The set of spectra is then scrunched over all frequency channels to obtain a time series. The time series median and standard deviation are calculated, and time samples lying over a threshold (e.g. 3 standard deviations from the median) are identified.

4. **Replace outliers**: Timesamples identified in Step 3 are replaced for each frequency channel. Samples are replaced with a random number drawn from a Gaussian with a mean equal to that of the time series divided by the total number of frequency channels, and a standard deviation of the channel[31].

The algorithm operates iteratively on relevant filterbanks within a given data directory[32]. As the algorithm is Python-based and mathematical operations are performed on 32-bit floats, output data products are may be scaled to unsigned 8-bit format if specified by a user[33]. Output clipped data products have amended `.rficlipped.fil` file extensions.

---

[30]Currently the clipping algorithm iteratively discards outliers which deviate > 3 standard deviations from the median.

[31]The standard deviation should ideally be unity due to rescaling in Step 2.

[32]Using: `/share/nas1/LOFT-e/software/analysis_pipeline/make_clipped_filterbanks_3.py`.

[33]During scaling, the mean and standard deviation of loaded spectra are calculated. Data is mean subtracted, divided by the standard deviation, and samples beyond 4 standard deviations from the mean are clipped.

Figure 5.10 shows the application of this algorithm on LOFT-e data. The figure compares frequency averaged time series for each e-MERLIN telescope and multiple sub-bands prior to and post RFI clipping. The RFI clipped data show significant reductions in amplitude of periodic RFI spikes when compared to their unclipped counterparts in Knockin, Pickmere, Mark II and Cambridge's sub-bands. Defford (1414 MHz) and Darnhall (1414 MHz) sub-bands show minimal improvement. The effects of application of this algorithm on S/N recovery in pulsar observations will be discussed in Chapter 6. As the technique preferentially removes significant, narrow zero-DM pulses from time series data, it may potentially remove astrophysical single pulses with extremely low DMs. Further investigation is necessary to estimate the impact it may have on detection completeness of low-DM events.

5.2.2.3   Frequency domain MAD flagging

The median absolute deviation (MAD) is a statistic used for identifying outlying samples in datasets. For a dataset x of N samples ($x_0, x_1, ...\ x_{N-1}$), the MAD is defined as:

$$\text{MAD} = \text{median}(|x_i - \text{median}(x)|). \tag{5.5}$$

Generally the standard deviation σ, defined as:

$$\sigma = \sqrt{\frac{1}{N} \sum_{i=0}^{N-1} (x_i - \langle x \rangle)^2}, \tag{5.6}$$

(where $\langle x \rangle$ is the dataset mean) is used to identify outliers, whereby data are excluded if they deviate from the mean by more than a certain multiple of the standard deviation. However if a dataset contains large values, these will be heavily weighted by the square term in the standard deviation compared with their deviation from the median. Therefore the MAD may be considered a more robust method for excluding outlying data points from a Gaussian distribution.

The MAD has been used to improve SNR measurements of pulsars by differentiating narrow pulses from Gaussian noise (e.g. by Cameron et al. 2017). During LOFT-e processing, the MAD is used to automatically identify and mask narrow sets of RFI-saturated frequency channels which arise sporadically and are thus unsuitable for the brute-force permanent masking masking as discussed in Section 5.2.2.1.
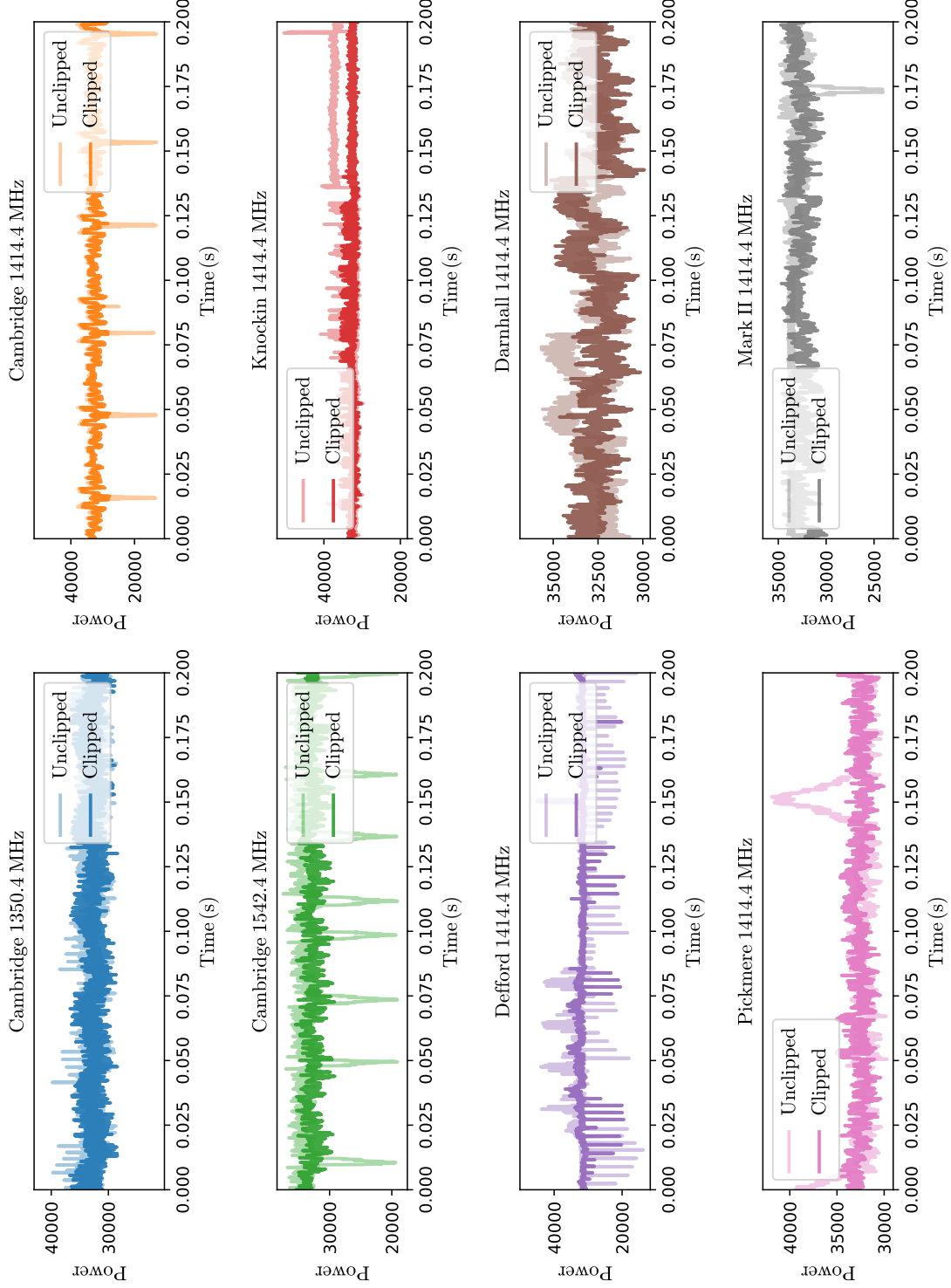
Figure 5.10: Timeseries data from different e-MERLIN telescopes and sub-bands, prior to and post-masking and RFI clipping. For each subplot frequency-averaged power is plotted against time into the observation. Timeseries plots were created by averaging standard LOFT-e output filterbank files (64-MHz-wide, 0.25 MHz spectral resolution, 64 μs temporal resolution). All data was recorded simultaneously.
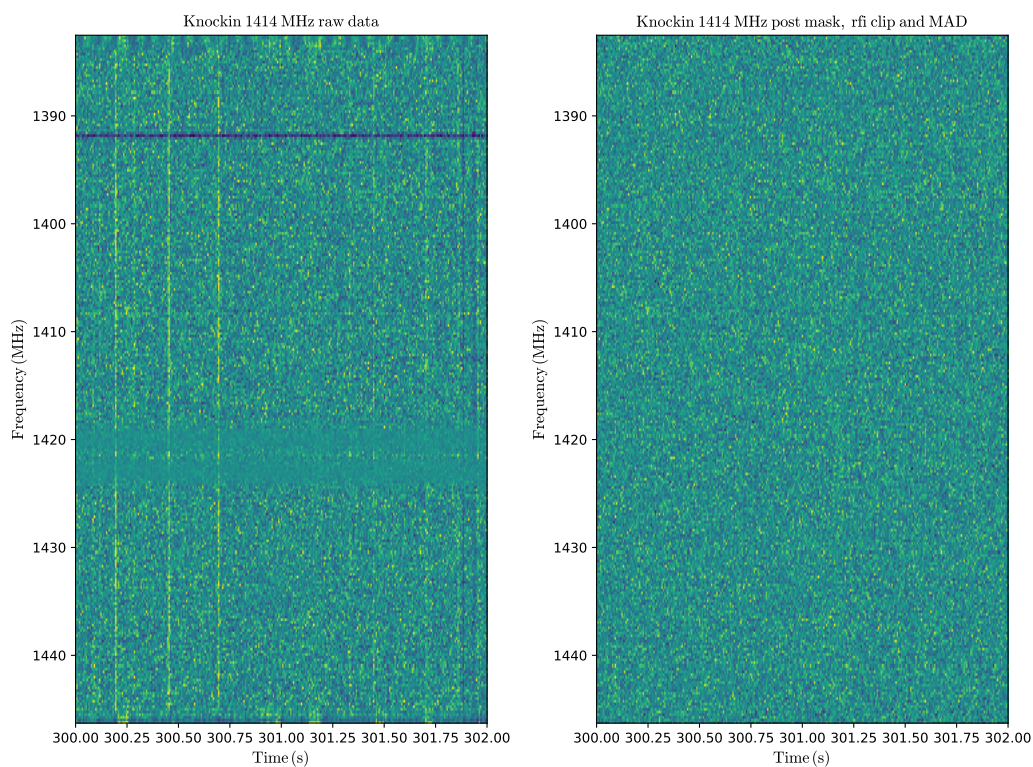
Figure 5.11: Comparison of e-MERLIN data acquired from the Knockin telescope prior to (left) and post-masking, time domain RFI clipping and frequency domain MAD flagging (right). This 64 MHz-wide sub-band centered on 1414.4 MHz was obtained using the Knockin telescope. Dynamic spectra resolution is as stated in Figure 5.8.

Frequency domain MAD flagging is performed by GPU-based code which operates faster than real-time[34]. During the flagging process, dynamic spectra are loaded spectra-by-spectra into GPU global memory, the median of each spectra is calculated via histogramming, and samples greater than a user-defined threshold from the MAD are masked[35]. MAD flagging is optionally applied to data files as the final stage of individual data stream-based RFI mitigation, post-RFI clipping and before data from telescopes and frequency sub-bands are combined[36]. The resulting flagged data products have amended `.MAD.fil` file extensions.

A comparison of a LOFT-e dynamic spectra prior to and after undergoing the full sequence of RFI mitigation techniques: masking, clipping, and MAD flagging is shown in Figure 5.11. The left-hand subplot shows persistent stripes of RFI at the band edges, impulsive broad-band vertical streaks of RFI and narrow-band RFI at ~ 1420 MHz and ~ 1392 MHz. In the right-hand subplot, this RFI has been removed.

Normalised histograms showing the histogrammed distributions of data after masking, time domain RFI clipping and frequency doman RFI flagging have been applied are presented in Figure 5.12. The data are processed versions of the same time series data sets presented in Figure 5.7 and may be compared on a telescope-by-telescope basis. In each case data takes on a visibly more Gaussian profile, however the multi-peaked structure of Knockin (1414 MHz), Defford (1414 MHz), and Cambridge (1414.4 MHz, 1542 MHz) has not been entirely removed. The effects of these techniques on pulse profile S/N recovery during pulsar observations is examined in Chapter 6.

5.2.3 Incoherent beamforming

In Chapter 3 it was emphasised that a $\sqrt{N}$ increase in sensitivity to fast radio transients may be achieved by an N-element array through by the incoherent combination of signals from the arrays individual elements (if the elements are identical). As radiometer theory relies on the assumption of Gaussian data distributions (Burke & Graham-Smith, 2002), the empirical sensitivity increase measured during incoherent beamforming may not reach theoretical predictions if non-Gaussian RFI is present in data.

---

[34]GPU MAD code developed by K. Rajwade and R. Joshi. Most recent version is located at: `/share/nas1/LOFT-e/software/MAD_pipeline/madfilter_small`.

[35]Currently samples > 3 MADs are replaced with the median value.

[36]By: `/share/nas1/LOFT-e/software/analysis_pipeline/make_MAD_filterbank_3.py`.
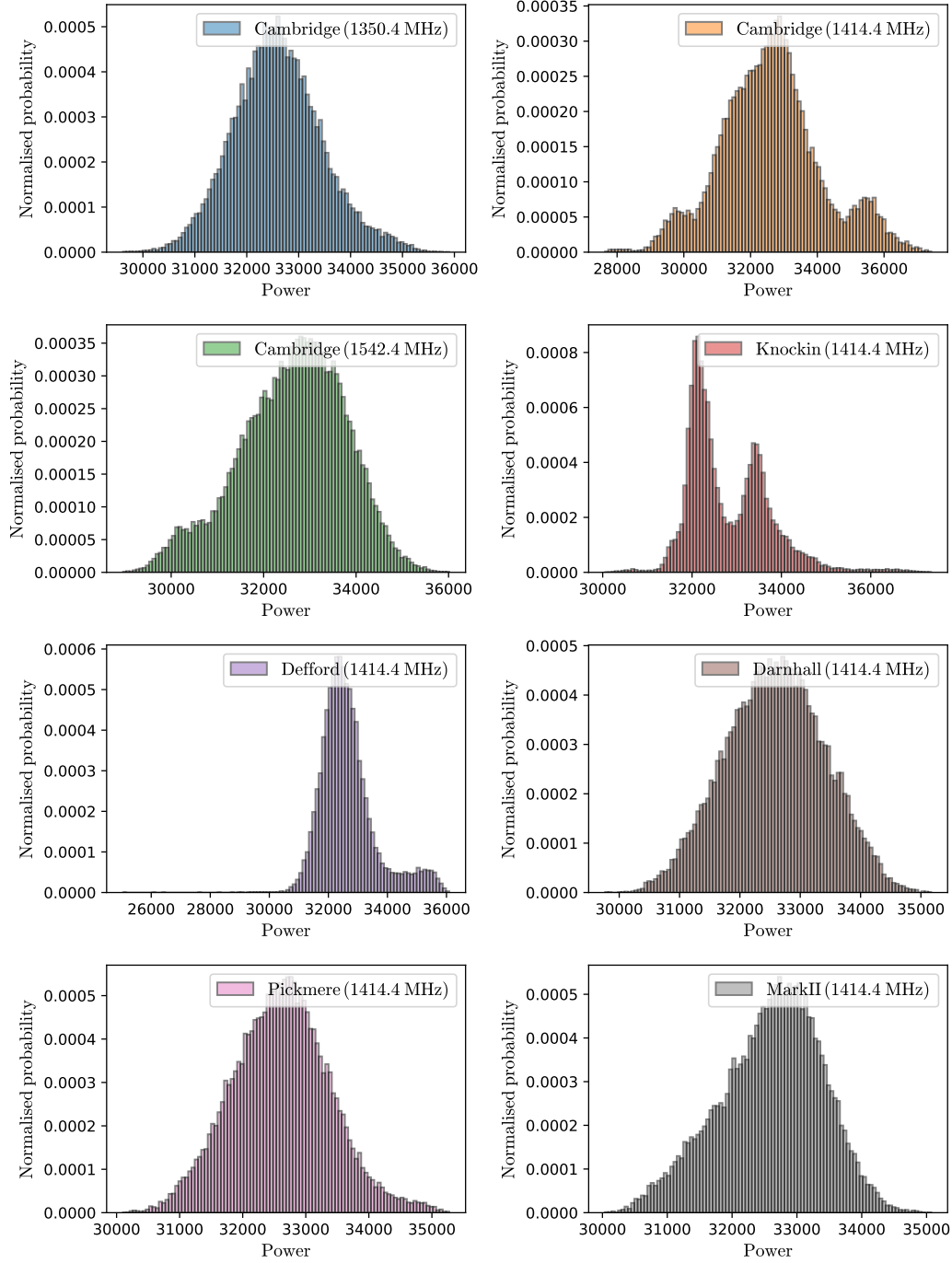
Figure 5.12: As Figure 5.7 for the same data sets after RFI mitigation techniques used in the LOFT-e data processing pipeline have been applied. These techniques consist of masking, time domain RFI clipping, and frequency domain MAD flagging.

As the incoherent sum of elements combines data from all elements of the array, the presence of bright RFI in a single element will increase noise levels beyond those predicted by Gaussian statistics. It may therefore be preferable to discard a data sample from an RFI-contaminated element, sacrificing a small signal increase[37] for a greater noise reduction in the presence of extreme RFI. Automated removal of RFI-contaminated samples during beamforming may be achieved by calculating the median value of signal recieved by each element on a sample-by-sample basis, and discarding the sample furthest from the median, prior to incoherent combination.[38] The theoretical maximum S/N increase for an N-element array using this technique scales as $\sqrt{N-1}$.

The impact of RFI on noise is compounded when present in multiple elements simultaneously. In this case, it may be less clear which, or how many samples are contaminated with RFI. If the RFI is terrestrial in origin however, the RFI signatures themselves are likely to vary between elements. Therefore by simply recording the median of the signals received by elements on a sample-by-sample basis, a more accurate representation of an RFI-free environment may be obtained[39]. The sensitivity increase afforded by this telescope combination method will not reach that of an incoherent beam, but will approach that of an RFI-free individual element. Depending on the severity of RFI this method may be preferable to incoherent beamforming for maximising an instrument's sensitivity.

During incoherent beamforming[40], LOFT-e may combine signals from individual dishes using all three of the above techniques. The resulting data combination modes will hereafter be referred to as *mean combination*, *median outlier rejection*, and *median combination* respectively. Algorithmic descriptions of these Python-based modes are described below.

1. **Mean combination** produces a standard incoherent beam from all available telescopes. A new filterbank file is initialised and a set of spectra is read for each telescope involved in the observation. The appropriate samples are incoherently summed, the output is scaled to 8-bit (if specified by the user)[41] and appended to the filterbank file. The next set of spectra is then loaded. Beamformed data

---

[37]A factor $\sqrt{6} - \sqrt{5}$ is lost for a standard six-dish, twelve baseline board LOFT-e configuration.

[38]The median, rather than the mean is chosen as the former is more robust to outliers.

[39]Again the median is taken rather than the mean as it is more robust to outliers.

[40]Most recent beamforming pipeline:
`/share/nas1/LOFT-e/software/incoherent_beam_pipeline/make_incoherent_beams_8.py`.

[41]Scaling specifications are the same as during RFI clipping.

generated using this mode contains the filename identifier `ibeam`.

2. **Median outlier rejection** (MOR) discards signals furthest from the median on a sample-by-sample basis. A new filterbank file is initialised, a set of spectra is read for each telescope, and median signals are calculated for each sample. The telescope with a signal deviating furthest from the median is then discarded. An incoherent sum of the remaining telescope signals is performed. The result is scaled to 8-bit (if specified) and appended to the filterbank file. The next set of spectra is then loaded. Beamformed data generated using this mode contains the filename identifier `mfbeam`.

3. **Median combination** takes the median of telescope signals for a given data sample. A new filterbank file is initialised, and a set of spectra is read for each telescope. For each sample, the median signal is calculated over all telescopes. The median signals are scaled to 8-bit (if specified) and appended to the filterbank file. The next set of spectra is then loaded. Beamformed data generated using this mode contains the filename identifier `mbeam`.

To retain storage space across LOFT-e's data-receiving hard drives, beamformed data products are stored in the 48 TB nas1 drive accessible by all three nodes.[42] Before beamforming begins, all relevant data and metadata are copied sequentially to LOFT-e's head node over a 1 Gb/s internal optical fibre[43]. Beamforming is performed sequentially for all sub-bands in a data directory automatically by utilising VDIF export metadata. Output data products are named for their observation MJD, sub-band center-frequency, file extension number, number of bits, and beamform mode[44].

Beamforming must be initiated on LOFT-e's head node, and as with earlier stages of data processing, is currently initiated via an end-user script[45]. All user inputs provided to the main processing pipeline must be provided, along with the desired beamforming mode. Examination of the sensitivity increase afforded by each mode during pulsar observations are found in Chapter 6.
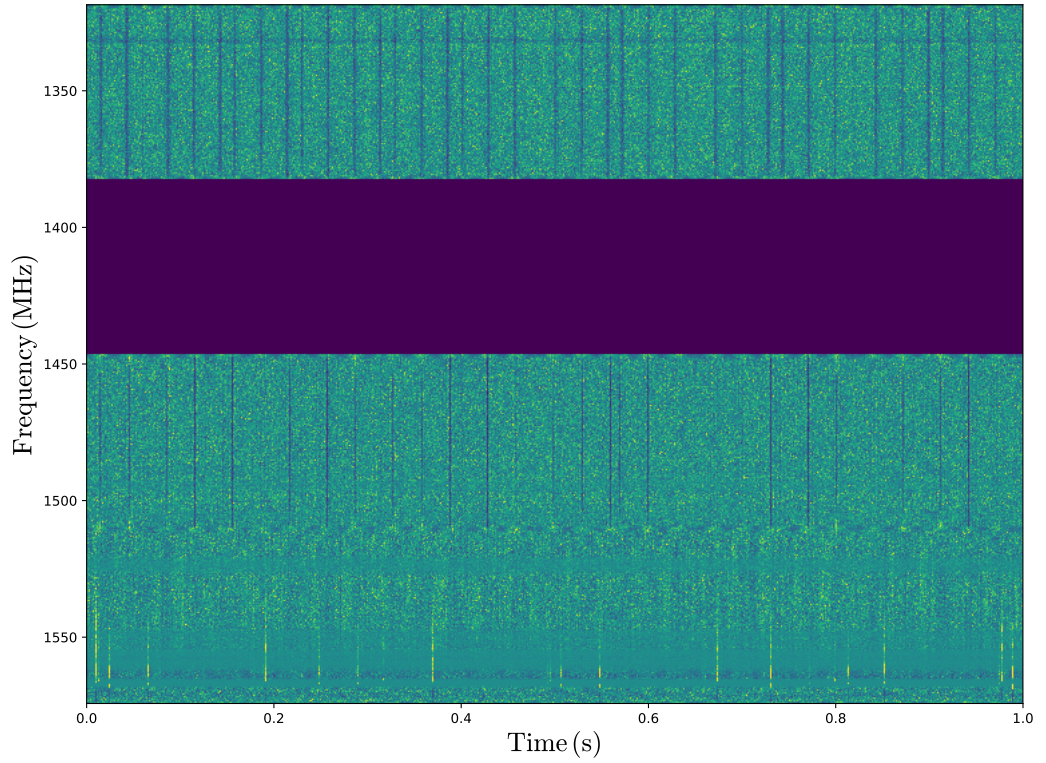
Figure 5.13: One-second dynamic spectrum of observing frequency against time formed from three 64-MHz-wide Cambridge frequency sub-bands observed simultaneously with LOFT-e. Each individual sub-band had 0.25 MHz spectral resolution and 64 μs temporal resolution. Colour denotes relative intensity. Channels where no data was recorded are filled with zero values and appear blue.

5.2.4   Sub-band combination

The final stage in LOFT-e's data processing pipeline is the combination of all frequency sub-bands recorded during observations. Theoretically, in the absence of RFI the sensitivity of an instrument to fast radio transients scales with the square root of bandwidth used according to the Equation 3.7. For the general six-telescope, twelve baseline board LOFT-e configuration, two sub-bands totaling 128 MHz bandwidth may be captured per telescope, affording a theoretical sensitivity increase of $\sqrt{2}$.

The data from certain e-MERLIN telescopes may occasionally be unsuitable or unavailable for use (e.g., if the telescope is off-source or under maintenance). With fewer telescopes, the lost sensitivity may be recovered by reallocating available baseline boards to increase captured bandwidth for those remaining. However if the baseline is shortened due to telescope outages, the angular resolution of the array will be reduced and its ability to localise transients will be compromised.

Frequency sub-bands captured by LOFT-e may be contiguous or non-contiguous. Selection of sub-bands is at a LOFT-e user's discretion, however the RFI environments of individual sub-bands should be considered when deciding on an experimental configuration. Early studies indicate that certain RFI-contaminated sub-bands may actively reduce the S/N of pulse profiles when included in pulsar observations. This will be discussed in Chapter 6[46]. Therefore sub-band combination is not currently an automated part of LOFT-e's data processing pipeline, and is performed manually by a user[47].

When non-continuous frequency sub-bands are combined, sub-bands empty intervening frequency channels are filled with zero values in LOFT-e filterbank data products. An example LOFT-e multiple sub-band filterbank file is shown in Figure 5.13. No RFI mitigation has been applied to these dynamic spectra. Periodic RFI is seen in all three sub-bands as dark vertical striping. As discussed in Section 4.5, this broadband impulsive RFI appears to occur at the same time in each sub-band. However, the RFI occurs with different intensities and is spread over different timescales in each sub-band.

---

[42]Beamformed products are stored in: `/share/nas1/LOFT-e/software/incoherent_beam_location/`.

[43]1 GB/s fiber is on transient's eth(3) interface and the compute nodes' (eth2) interfaces. Data could be transferred faster over the 10 GB/s eth(0) interface instead by using correct IP adresses.

[44]E.g.: `mjd_58059.7320752_ibeamCF_1414.4MHz.File1.8bit.fil`.

[45]E.g.: `/share/nas1/LOFT-e/software/incoherent_beam_pipeline/make_incoherentscript_180808.py`.

[46]A comprehensive study of e-MERLIN's RFI environments on a sub-band by sub-band basis at various observing times and source latitudes is currently a subject of research by T. W. Scragg.

[47]Sub-bands are combined using pulsarsoft's `filmerge` software.

5.3 LOFT-e transient analysis techniques

Output filterbanked data products generated by the LOFT-e pipeline may be searched for transient phenomena at several stages during data processing: post-initial filterbanking, post-beamforming, and post-sub-band combination (see Figure 5.5). Methods used to search for single pulses vs pulsars differ, so techniques generally applied during fast transient analysis will vary based on the scientific aims of an experiment. General overviews of single pulse and pulsar detection are provided in this Section, and placed in context of LOFT-e's current capabilities.

In the cases of both single pulse and pulsar detection, events are generally identified from frequency-scrunched time series products (see Section 4.2.2). Therefore to achieve maximum S/Ns for events detected in data, the sensitivity-reducing impact of dispersion (see Section 2.1) must be accounted for. This process is known as dedispersion. Dedispersion may either be performed incoherently or coherently (see, e.g., Cordes & McLaughlin (2003) which contains an overview of both methods in the context of fast transient detection).

Incoherent dedispersion techniques neglect the phase information of data and are less computationally expensive, while coherent dedispersion techniques allow for improved time resolution and thus the study of events fine temporal structure (see, e.g. Farah et al. 2018). Currently, incoherent dedispersion is used in LOFT-e's pipelines and so will be overviewed in this Section. Due to LOFT-e's capability to capture voltage data, coherent dedispersion techniques may be implemented in the future (see Section 5.4).

5.3.1 Incoherent dedispersion

During the process of incoherent dedispersion, time shifts are applied to samples in a filterbanked dynamic spectrum in order to reverse the frequency-time dependent delay measured by particular DM. The act of incoherent dedispersion forms a dedispersed filterbank data product, which may then be averaged over frequency channels generating a dedispersed time series. This time series may then be searched for transients. As per the notation established in Section 4.2.2, a dedispersed time series created from the dynamic spectrum $D_s(t_i, F_j)$ of fourier components F_j for time samples t_i may be mathematically described as:
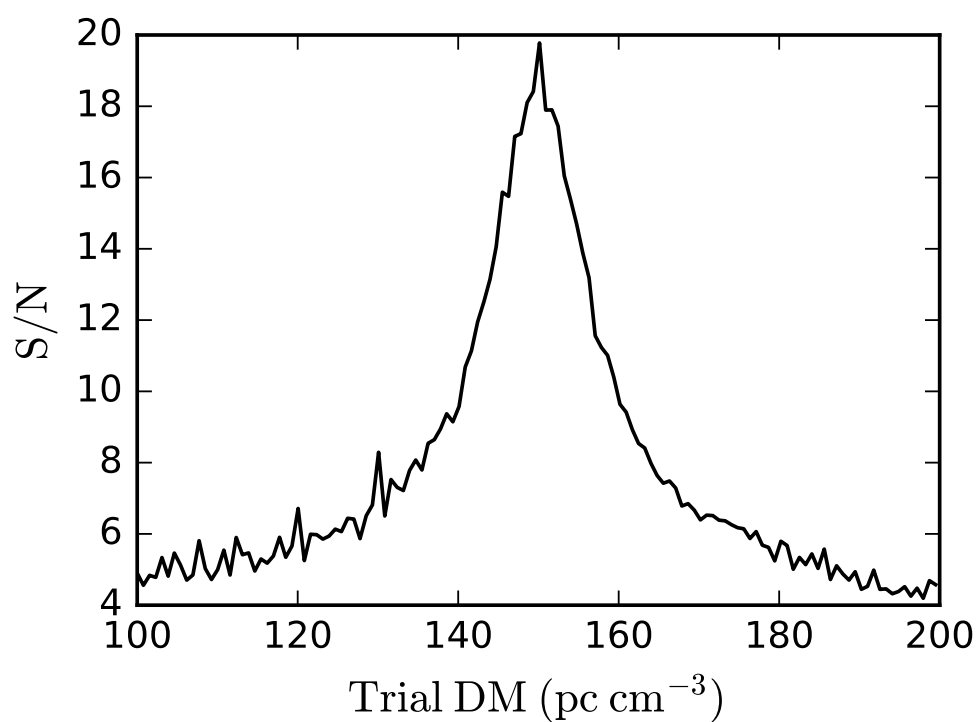
Figure 5.14: S/N of the pulse from Figure 2.1 recovered by incoherent dedispersion at a trial DM DM_{trial} for many trial DMs between 100 pc cm^{-3} and 200 pc cm^{-3}. At the DM_{trial} closest to the pulse's true DM value, maximum S/N is recovered.

$$T_i = \sum_{j=1}^{n_{chan}} D_s(t_i - \Delta t(\nu_j), \nu_j), \tag{5.7}$$

where ν_j is the frequency corresponding to Fourier component F_j, and the frequency-dependent time shift $\Delta t(\nu_j)$ is calculated using Equation 2.3 by substituting ν_1 and ν_2 for ν_{ref} and ν_j respectively. Here the reference frequency ν_{ref} is the center of the highest channel of the filterbank file.

When optimally dedispersed the maximum S/N of a transient event will be recovered. Figure 5.14 illustrates this effect, showing the resulting S/Ns recovered for a simulated pulse with intrinsic DM 150 pc cm^{-3} after incoherent dedispersion to many different 'trial' DMs. When studying a known source with a well-measured DM (e.g. during routine pulsar timing), data may be dedispersed with the correct DM immediately. However when searching for new transients, the DMs of any sources which happen to arise in observations are not already known. Therefore during searches a single data set may be dedispersed to a significant number of trial DMs.

As the DM of a fast radio transient depends on its distance and the distribution of electrons along its sightline the DM values of known events vary wildly. Pulsars have DMs between ~ 2 pc cm^{-3} (PSR J1744$-$3933, Tauris et al. 1994) and ~ 1700 pc cm^{-3} (PSR J1447$-$2900, the magnetar lying close to the Galactic center Mori et al. 2013). At the time of writing, the most extreme DM measured for an FRB is ~ 2600 pc cm^{-3} (FRB 160102, Bhandari et al. 2018). Therefore the number of trial DMs searched and the choice of DM step size between trials must be optimized. Too small a step size will result in the creation of multiple near-identical dedispersed time series, wasting computing resources. Too large a step size, and potential events in observations with DMs between the trials may not be sufficiently dedispersed to reach a threshold S/N signifying a detection.

As reviewed by Lorimer & Kramer (2012), a pulse of intrinsic width W_{int} dedispersed using a DM offset from its true DM by ΔDM will be smeared in time across its entire observing bandwidth $\Delta\nu$ by W_{DM}, such that $W = \sqrt{W_{int}^2 + W_{DM}^2}$ where:

$$W_{DM} = \left(\frac{8.3 \times 10^6}{ms}\right)\left(\frac{|\Delta DM|}{pc\ cm^{-3}}\right)\left(\frac{\Delta\nu/\nu^3}{MHz}\right), \tag{5.8}$$

when the observing centre frequency $\nu >> \Delta\nu$. A pulse in a filterbanked data product with sampling interval t_{samp} will be sufficiently dedispersed when $W_{DM} = t_{samp}$. For a given observation, the ΔDM associated with this smearing is therefore often chosen as the step size for DM trials. Each trial DM has its own frequency-dependent dispersive

delay. When a sufficiently large DM, the 'diagonal DM', is reached that smearing in a single channel exceeds t_{samp}, adjacent time samples are combined to preserve computing resources (Lorimer & Kramer, 2012). Algorithms with differing computational costs (see, e.g., The CHIME/FRB Collaboration et al. 2018 for a brief overview) exist to calculate the optimum DM step routine for transient search pipelines.

5.3.2 Pulsar analysis

Following dedispersion, the resulting time series are searched for transients. New periodic sources (e.g. pulsars) are generally identified by periodicity searches[48]. During analysis of known or new pulsars (e.g. study of their pulse profiles) a single pulse may not be visible above the background noise of observations. Therefore time series data may be folded at the period of the pulsar such that pulses are phase-aligned. The folded data may then be summed to form an integrated pulse profile. Ephemerides used to account for effects such as pulsar spin-down rate, sky position, proper motion and a pulsar's motion if in a binary system may be utilised.

Folding results in a 3-dimensional array of time samples (bins), frequency channels (sub-bands) and subintegrations. The center of each bin of a subintegration corresponds to a certain point (phase) in one pulse period. Folded profiles may be averaged over subintegrations to form pulse profiles, or examined as a function of pulse phase vs time, or pulse phase vs frequency, allowing the study of effects such as scintillation.

During LOFT-e analysis[49], folding is performed according to an ephemeris file provided by the user upon running the pipeline[50]. Folding is performed using the DSPSR pulsar data processing package[51] and resulting data products are output as archive files[52]. Output products are named according to their input filterbank files with an appended `.ar` file extensions. Folded pulse profiles, bandpasses, and pulse phase vs frequency/time plots are also automatically created for visual examination. The

---

[48]During a periodicity search, an FFT is performed over the dedispersed time series and the power of resulting fourier components ($|F_k|^2$) are plotted against their corresponding frequencies ν_k and examined for significant features. See, e.g., Lorimer & Kramer 2012 for an overview. Currently LOFT-e does not implement periodicity searches during transient analysis.

[49]Initiated by `/share/nas1/LOFT-e/software/analysis_pipeline/fold_fil_list_1.py`.

[50]Ephemeres may be obtained using, e.g., the pulsar catalogue (`http://www.atnf.csiro.au/research/pulsar/psrcat/`) (Accessed: 25/09/2018) and currently must be stored in `/share/nas1/LOFT-e/ephemeris/`.

[51]`http://dspsr.sourceforge.net/` (Accessed: 25/09/2018) (van Straten & Bailes, 2011)

[52]Other common packages used for folding/pulsar searching PRESTO (`https://www.cv.nrao.edu/~sransom/presto/` (Accessed: 25/09/2018)) and SIGPYPROC (`https://github.com/ewanbarr/sigpyproc`).
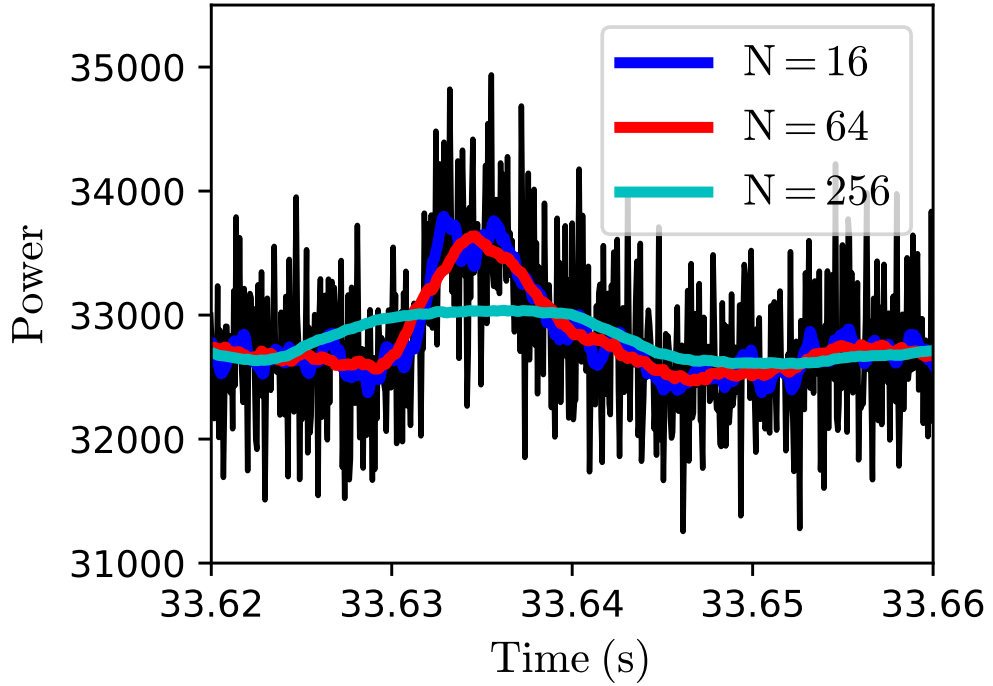
Figure 5.15: Dedispersed, frequency-scrunched power against time from the dynamic spectrum acquired during a LOFT-e observation of RRAT J1819+1458. A single pulse occurs at ~ 33.635 s into the observation. Overplotted are data at native filterbank temporal resolution ($t_{\text{samp}} = 64\,\mu$s (black line) and data filtered with boxcars of width $N \times t_{\text{samp}}$.

analysis of pulsars observed by LOFT-e are discussed in Chapter 6.

5.3.3 Single pulse analysis

Fast radio transient phenomena such as pulsar giant pulses, RRAT bursts, and FRBs are detected in a dedispersed time series using single pulse search techniques. Such techniques are designed to identify signals which exceed a user-defined threshold above the noise in an observation (see, e.g., Cordes & McLaughlin 2003; Lorimer & Kramer 2012 for reviews of this topic)..

Historically, transients discovered with single pulse searching have been found during post-analysis of archival pulsar survey data. The Lorimer burst (Lorimer et al., 2007) and first RRAT detections (McLaughlin et al., 2006) were discovered in this way, as were many other FRBs (e.g. Thornton et al. 2013). Increasingly, however,

modern pipelines implementing GPU-based dedispersion and single pulse search techniques[53] are able to search for single pulses in real time and operate commensally to standard observations. This is beginning to expand the scope of fast transient research beyond centimetre wavelengths, and even beyond the electromagnetic spectrum. For example the SUPERB collaboration has performed rapid multi-messenger follow-up with optical, X-ray, gamma-ray and neutrino observatories while searching for counterparts to FRBs (Keane et al., 2014; Bhandari et al., 2018).

LOFT-e currently uses the GPU-based AstroAccelerate[54] software for dedispersion and single pulse searching. The default DM range searched is 0-3000 pc cm^{-3}. This range contains the DMs of all catalogued FRBs to date[55] and the DM contribution of most lines of sight in the Milky Way (Figure 2.2; Cordes & Lazio 2002).

AstroAccelerate employs matched filtering to detect single pulses. During matched filtering, a dedispersed time series $T = \{t_i\}_{i=1}^{N}$ of N time samples is convolved with a kernel of a specified width n. In the simple case employed by AstroAccelerate, the kernel is a boxcar[56], and the result of convolution is a new time series $T_{\mathrm{av}} = \{t_i'\}_{i=1}^{N-n+1}$ which is the n-moving average of T[57] :

$$t_i' = \frac{1}{n} \sum_{j=i}^{i+n-1} t_j.$$
(5.9)

Boxcar filtering of a pulse-containing time series with a window wider than the pulse width will result in noise being averaged with signal and a less than optimal S/N will be recovered. Utilising a window smaller than the pulse width will result in some signal being excluded. Therefore single pulse search algorithms search time series filtered over many different boxcar widths for signals over a detection threshold (Lorimer & Kramer, 2012). Figure 5.15 demonstrates the effect of boxcar filtering with multiple different windows on a single pulse from RRAT 1819-1458, found in a LOFT-e filterbanked data product with $t_{\mathrm{samp}} = 64\,\mu$s time resolution. The time series filtered over a window of $n = 64 \times t_{\mathrm{samp}}$ provides the SNR closest to the AstroAccelerate-recovered pulse width of 4 ms.

---

[53]GPU-based software includes AstroAccelerate (`https://github.com/AstroAccelerateOrg/astro-accelerate`) (Adámek & Armour, 2016) and Heimdall (`https://sourceforge.net/p/heimdall-astro/wiki/Use/`) (Accessed: 25/09/2018).

[54]`https://github.com/AstroAccelerateOrg/astro-accelerate/wiki`

[55]`http://frbcat.org/`; (Petroff et al., 2016)

[56]More complicated kernels may be employed by weighting time samples to match a template pulse.

[57]Weisstein, Eric W. "Moving Average." From MathWorld–A Wolfram Web Resource. `http://mathworld.wolfram.com/MovingAverage.html` (Accessed: 25/09/2018).

LOFT-e single pulse searching is currently initialised by an end-user script[58] within which the threshold S/N for discarding candidates[59], the maximum boxcar window[60] and directories containing data to be searched are specified. Following initialisation, parameter files necessary for AstroAccelerate are generated automatically and files are searched sequentially for single pulses[61]. Candidate files[62] may then be plotted[63] and examined by eye.

5.3.3.1 Zero DM RFI mitigation

In addition to the RFI mitigation techniques discussed in Section 5.2.2, a further RFI mitigation technique is built into AstroAccelerate and applied during LOFT-e's single pulse searching. Zero DM RFI mitigation exploits the fact that RFI has not propagated through ionised electrons, and so in general will not display a frequency-dependent dispersive sweep (such as seen in Figure 2.1). As such it will have a maximum S/N at DM=0 pc cm^{-3}, as opposed to astrophysical phenomena which will be dispersed over the observing band[64].

To apply this technique to a dynamic spectrum, the Fourier components of each frequency channel are averaged for each spectra in the file, resulting in a time series of mean values. Each Fourier component in the dynamic spectrum is then divided by its respective mean value, reducing the prominence of zero DM RFI. The process can be useful for reducing the false candidate detections generated during searching. However, the technique can modulate the intensity of dispersed pulses across the observing bandwidth, and so should not be applied when studying, e.g., the scintillation properties of fast transient events[65].

---

[58]E.g.: `/share/nas1/LOFT-e/software/single_pulse_pipeline/run_single_pulse_pipeline.py`.
[59]E.g. 8σ.
[60]Which should be wider than the probable width of expected pulses.
[61]Pipeline software:
`/share/nas1/LOFT-e/software/single_pulse_pipeline/single_pulse_pipeline_1.py`.
[62]Stored in: `/home/cwalker/astro-accelerate_charlie/astro-accelerate/output/`.
[63]With:
`/share/nas1/LOFT-e/software/single_pulse_pipeline/plot_software/plotEventsScript.py`.
[64]An exception to this rule are *Perytons*, whose signature appears similar to a dispersively swept pulse (Burke-Spolaor et al., 2011).
[65]As the technique is used during LOFT-e's single pulse search process, but is not permanently applied to data products, this is not an issue during candidate analysis.

5.4   Future improvements to LOFT-e pipelines

LOFT-e's current data processing pipeline and techniques incorporated to mitigate e-MERLIN's RFI environment have been reviewed in this chapter. The pipeline has successfully been used to detect known pulsars and RRAT single pulses during test observations. Analysis of the improvements afforded by the stages of the pipeline on such observations will be discussed in Chapter 6. However in order to maximise LOFT-e's potential as a fast transient facility, several improvements should be implemented into future pipeline iterations. These include:

1. **Full commensality**: LOFT-e's current pipeline operates on data stored to LOFT-e's disks. Due to LOFT-e's low expected FRB detection rates, a full commensal pipeline running in parallel to regular observations will be necessary for LOFT-e to detect events. Development of the noncommensal pipeline has been used as a testbed for LOFT-e's eventual real-time commensal mode, allowing challenges specific to e-MERLIN (e.g. instrumental, RFI, and data-flow challenges) to be identified. In LOFT-e's commensal mode, data will be filterbanked, processed, dedispersed, and searched for FRBs in real time, and raw data surrounding candidates will be stored for offline localisation. For this to be achieved, LOFT-e's Python-based time-domain RFI clipping and beamforming software must be updated to operate faster than real time. When full commensality is achieved, the ability to alert the FRB research community to events detected by e-MERLIN may be implemented. Standardising observatories' ability to notify the transient community potential FRB candidates may come from use of the Virtual Observatory Events (VOEvents) framework (Petroff et al., 2017).

2. **Output of all Stokes data products**: During filterbanking, LOFT-e currently squares and combines e-MERLIN's two polarisations to output Stokes I (total intensity) data products. Future modifications to the filterbanking software must be made to allow further Stokes parameters to be calculated in order for polarisation properties of LOFT-e detected events to be analysed.

3. **Coherent beamforming capability**: As discussed in Chapter 2, coherent beamforming affords an additional $\sqrt{N}$ increase in sensitivity over incoherent beamforming for an N-element interferometer. By observing phase calibrators prior to pulsar observations and capturing this data, LOFT-e may be able to calculate the phase delays necessary for LOFT-e to form coherent beams, theoretically increasing its sensitivity by up to a further factor of $\sqrt{6}$, provided 6 telescopes

are used during observations. Work has begun adapting Giant Metrewave Radio Telescope (GMRT) coherent beamforming software for this purpose.

4. **Coherent dedispersion**: Coherent dedispersion involves convolving voltage data acquired from a telescope with the inverse of a so-called transfer function. The transfer function describes how a radio signal propagating through an ionised medium is modified in terms of its phase component. As such, coherent dedispersion is applied to raw voltage data as opposed to squared filterbank data products. The coherent dedispersion process is reviewed in detail in Lorimer & Kramer (2012) and Bassa et al. (2017a). Coherent dedispersion has the advantage of completely removing the DM smearing of a transient event, as opposed to incoherent dedispersion which does not account for broadening within the individual frequency channels of a dynamic spectrum. This has been shown to improve detection of transients at lower observing frequencies (Bassa et al., 2017a) and results in improved time resolution data products and more accurate pulse profiles for detected fast radio transients (Lorimer & Kramer, 2012). As discussed by The CHIME/FRB Collaboration et al. (2018), the fine structural details of FRBs attainable via coherent dedispersion (see, e.g., Farah et al. 2018) may provide insight into their emission mechanisms. LOFT-e's ability to capture voltage data enables potential coherent dedispersion of any fast radio transients detected with its pipelines. Modern GPU techniques are allowing coherent dedispersion algorithms to be efficiently implemented into transient search pipelines (Bassa et al., 2017a).

Chapter 6

LOFT-e: Experiments and Results

This chapter builds upon the introductory chapters introducing the study of pulsars and fast radio transient phenomena and the development of a high time-resolution backend to detect them with the e-MERLIN interferometer, by presenting observational results obtained by LOFT-e. The first successful pulsar observations with LOFT-e are presented, along with examples of observed behaviour such as mode switching, and propagation effects such as scintillation (which were reviewed in Chapter 2). The effects of LOFT-e's data processing stages (e.g. RFI mitigation, sub-band combination and incoherent beamforming, as introduced in Chapter 5) on sensitivity to sources are demonstrated and discussed. The first LOFT-e detection of a single pulse from an RRAT is presented, and future work to localise Galactic RRATs is overviewed. Finally, the first stage of an investigation into the sensitivity losses suffered by e-MERLIN's telescopes is reviewed.

6.1   Pulsar detection with LOFT-e

This section discusses early efforts to observe pulsars with LOFT-e and presents the first successful observations of pulsars using LOFT-e. LOFT-e's data processing stages, including RFI mitigation, sub-band combination and incoherent beamforming are demonstrated and LOFT-e's current sensitivity to pulsars is discussed. Parameters for all pulsars observed in this Section are provided in Table 6.1. Unless otherwise stated, peak S/Ns of pulse profiles presented in this section are calculated using:

$$S/N = \frac{I_{peak}}{RMS\ noise},$$

(6.1)

Figure 6.1: Observed Stokes I power vs. time for data captured during early LOFT-e C-band observations with e-MERLIN's Knockin, Darnhall and Mark II elements. Continuous AGC scaling was applied by the correlator during observations. Timeseries data was created from 200-channel dynamic spectra centred at 4975.84 MHz with an 250 μs integration period and 0.32 MHz spectral resolution.

where I_{peak} is the peak intensity of the frequency-averaged folded timeseries, and the RMS noise is calculated after masking of the pulsar's signal from the folded profile.

6.1.1   Unsuccessful pulsar observations and e-MERLIN's AGC

During early LOFT-e trial observations of bright pulsars, pulses were not routinely recovered. By analysing timeseries data created for individual telescopes and sub-bands for LOFT-e data products, non-Gaussian including periodicities, dips in power and varying dynamic range were detected, examples of which are presented in Figure 6.1. This figure compares 150 seconds of timeseries data acquired for three telescopes during C-band observations of PSR B0329+54. Observations took place on 13/12/2016 using e-MERLIN's Knockin, Darnhall and Mark II elements.

Data for each telescope display differing behaviour. Knockin and Mark II data show periodic signals. The Mark II's observed power additionally shows a long timescale drift, rising significantly over the course of the observation.

To constrain potential sources of observed non-Gaussian behaviour in LOFT-e data,
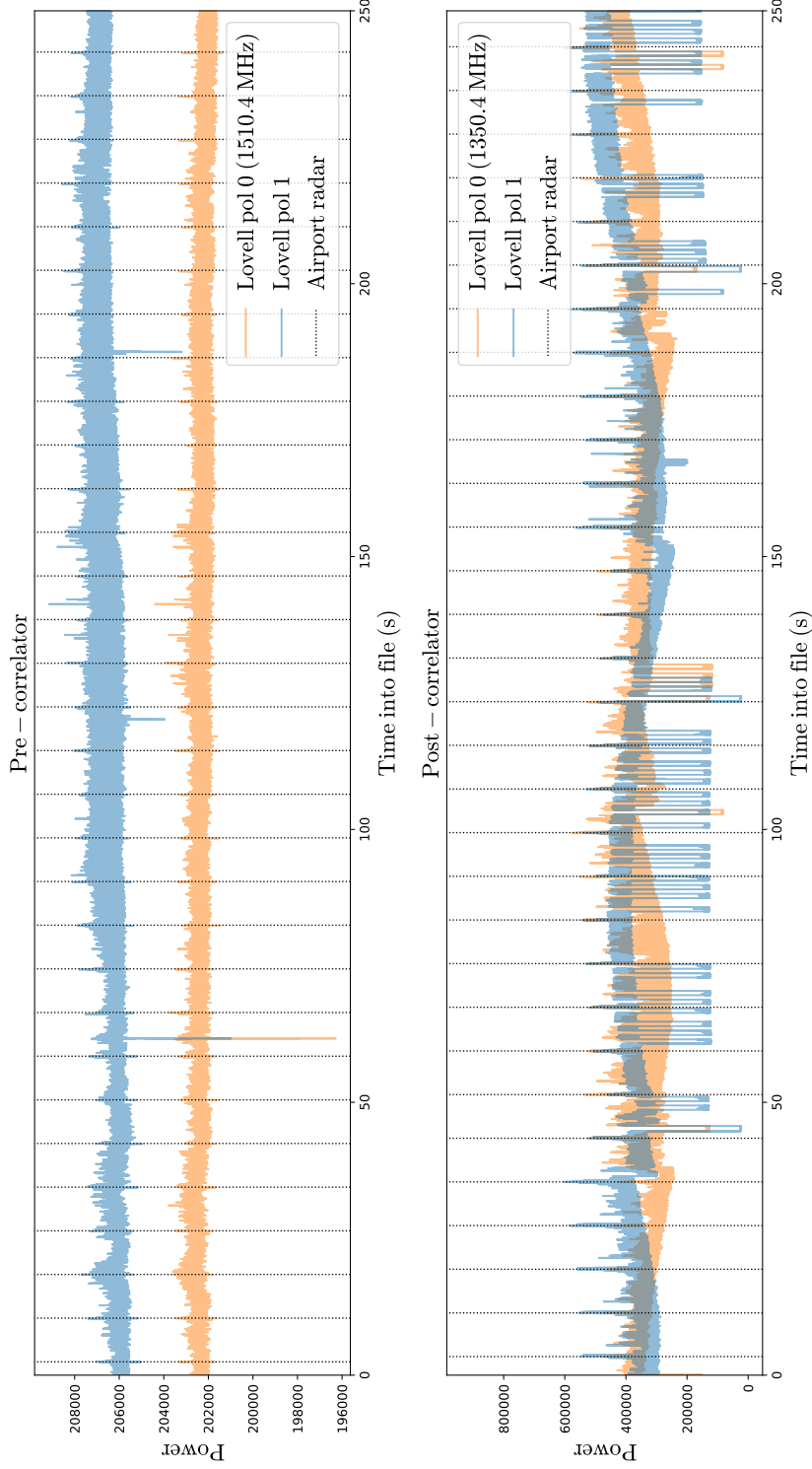
Figure 6.2: Observed power *vs.* time for observation of PSR B1933+16 with the Lovell Telescope. Data was captured prior to entering the e-MERLIN correlator (top) and by LOFT-e (bottom). Continuous AGC scaling was applied by the correlator during observations. For each data capture method, individual polarisations (upper subplot) are shown in solid colours. Vertical dotted lines indicate RFI spikes associated with airport radar. Pre-correlator timeseries data was created from 4096-channel dynamic spectra centred at 1510.4 MHz with 1024 μs temporal resolution and 0.125 MHz spectral resolution. Post-correlator timeseries data was created from 200-channel dynamic spectra centred at 1350.4 MHz with 250 μs temporal resolution and 0.32 MHz spectral resolution.
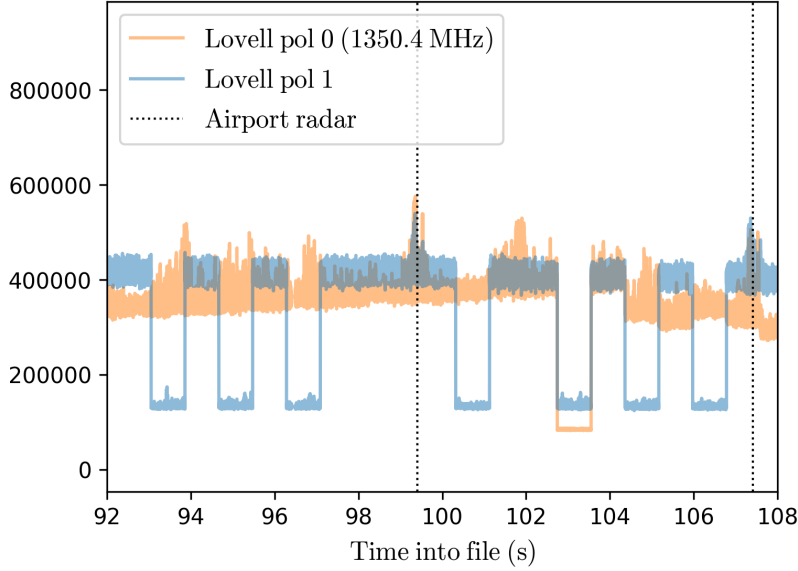
Figure 6.3: Close-up of observed power vs. time for observation of PSR J1933+16 with the Lovell Telescope. Data was captured using LOFT-e. The figure style and data parameters follow Figure 6.2 (second subplot). Both polarisations are plotted separately. Vertical dotted lines at ~ 99 s and ~ 107 s mark RFI spikes consistent with airport radar with an 8 s period.

an experiment using the Lovell Telescope was devised. A ~ 10 minute observation of PSR B1933+16 was recorded with an early version of the LOFT-e data capture pipeline. Within this observation period ~ 4 minutes of Lovell Telescope data were recorded prior to entering the correlator[1]. Timeseries products were created for each version of Lovell data and compared. These are presented in Figure 6.2. Pre-correlator data (top) was created from 4096-channel dynamic spectra centred at 1510.4 MHz with 1024 μs temporal resolution and 0.125 MHz spectral resolution and demonstrates the behaviour of the entire 512 MHz observing band. As only 64 MHz bandwidth was received by LOFT-e, post-correlator timeseries data (bottom) was created from 200-channel dynamic spectra centred at 1350.4 MHz with 250 μs temporal resolution and 0.32 MHz spectral resolution. For both data sets, individual polarisations are presented. The 250-second pre-correlator timeseries was captured during the ~ 600-second LOFT-e observation period.

Both pre and post-correlator data show RFI spikes with an 8-second period which are marked by vertical dashed lines in each subplot. These are associated with local

---

[1]This was achieved using a parallel signal branched out upstream of the e-MERLIN correlator utilising the Lovell Spectrometer (D'Cruze, 2018).

Source Name	P	$\dot{P}$	DM	RA	Dec
	s	10^{-15}s/s	pc/cm^{-3}	hh:mm:ss.s	dd:mm:ss.s
PSR B1933+16	0.359	6.00	158.5	19:35:47.83	+16:16:39.99
PSR B0329+54	0.715	2.04	26.7	03:32:59.37	+54:34:43.57
RRAT J1819−1458	4.263	575.50	196.0	18:19:34.17	−14:58:03.57

Table 6.1: Rotation, dispersion and position parameters for all pulsars and RRATs observed by LOFT-e and with data presented in this work. Parameters are obtained from (Hobbs et al., 2004; McLaughlin et al., 2006; Keane et al., 2011) and were supplied by the pulsar catalog (Manchester et al., 2005) and the RRATalog (`http://astro.phys.wvu.edu/rratalog/`) on 24/09/2018.

airport radar. LOFT-e data additionally exhibits a variety of non-Gaussian behaviours which are not visible in pre-correlator data. A long timescale drift in power (similar to that seen in Figure 6.1 Mark II data) is seen in individual polarisations.

Figure 6.3 shows a close-up view of LOFT-e captured data. At this resolution, periods of significantly decreased power corresponding to decreased dynamic range are visible. Such behaviour does not occur in data captured prior to entering the correlator, and so must be induced within the correlator. During these experiments, the AGC was applied continuously to data prior to its output from the correlator's baseline boards. Periodic rescaling by e-MERLIN's AGC was suspected as a source of non-Gaussian behaviour in data which was affecting pulsar recoverability.

When a command-line call from LOFT-e's head computing node is called, e-MERLIN's AGC may be activated manually, calculating scaling factors from a few seconds of data and applying them to all further data until the AGC is re-triggered. During later LOFT-e observations, the data capture pipeline was amended to utilise this call in an automated manner. This triggering was set to occur at the beginning of new observations (i.e. when telescope metadata indicated a change in source). In this way scaling factors would ideally be calculated semi-regularly, but would not take place mid-observation.

6.1.2 Successful pulsar observations and RFI mitigation

The new method of AGC application allowed the creation of visible pulse profiles from LOFT-e data products during pulsar observations. Timeseries data aquired using manual AGC triggering is presented in Figure 6.4. Knockin, Darnhall and Mark II data are again shown, allowing for comparison with Figure 6.1. Each telescope now

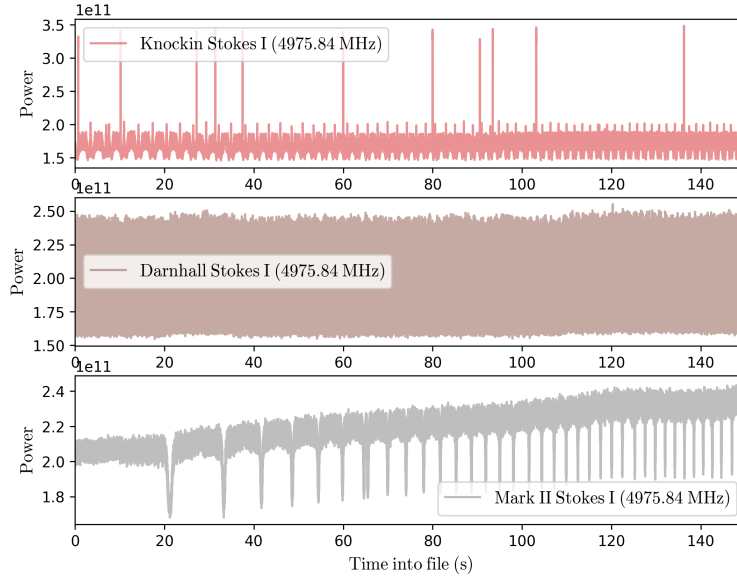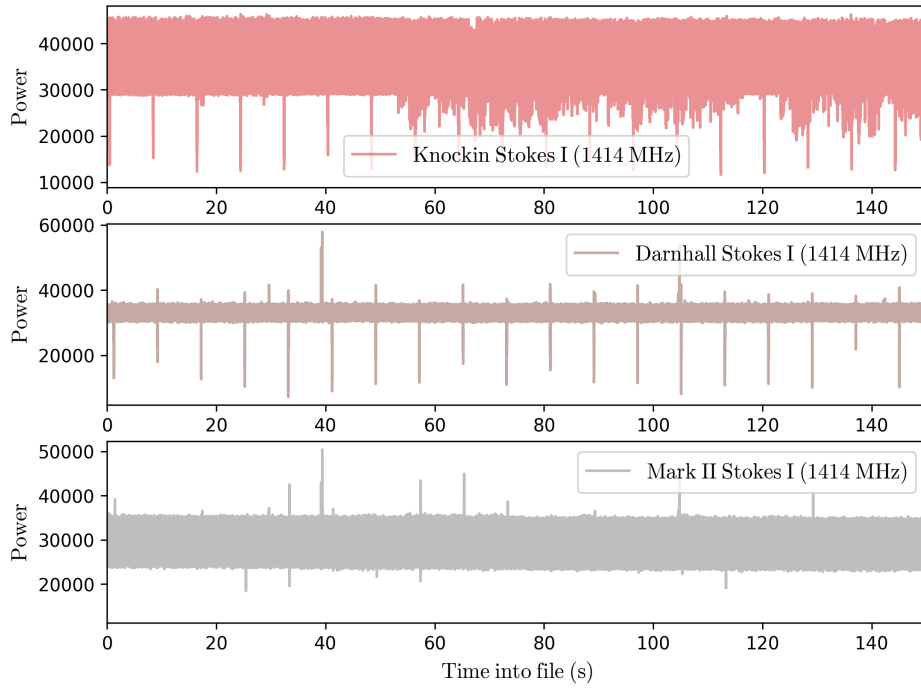**Figure 6.4:** Observed Stokes I power vs. time for data captured during LOFT-e L-band observations with e-MERLIN's Knockin, Darnhall and Mark II elements. The correlator AGC was triggered prior to the observation. Timeseries data was created from 256-channel dynamic spectra centred at 1414.4 MHz with an 64 μs integration period and 0.25 MHz spectral resolution.

shows a stable power level over time, indicating an improvement over continuously AGC-scaled data. It is noted that observing centre frequencies were not consistent between these observations. LOFT-e observations are subject to technical constraints such as available telescopes, installed receiver, and number of functional baseline boards (which impact available sub-bands). Therefore comprehensive analysis of the behaviour of data for all permutations of telescope, baseline-board and observing band has not yet taken place.

In successful pulsar observations, different telescopes and their frequency subbands have shown varying degrees of time-dependent RFI which may obscure their pulse profiles. The RFI mitigation techniques overviewed in Chapter 5 are often necessary to improve the sensitivity of such data. Results of RFI mitigation applied to pulsar observations at multiple sub-bands for e-MERLIN's telescopes are presented below.

6.1.2.1   PSR B0329+54 observations at multiple sub-bands

Figure 6.5 presents intensity against pulse phase plots created for PSR B0329+54 data acquired by LOFT-e at L-band. Each subplot shows data captured from a different 64 MHz-wide sub-band. Lines of successively darker shades in each sub-plot indicate data after application of each successive RFI mitigation stage discussed in Chapter 5. Prior to RFI mitigation, PSR B0329+54's average pulse profile is visible at 1414.4 MHz and 1670.4 MHz. The profile's main component peaks at ~ 0.17 and ~ 0.21 pulse phase respectively. The pulse profile is obscured at 1350.4 MHz and 1542.4 MHz by severe noise. At all centre frequencies, the template masks which were applied were manually created. These masks removed the ~ 8 frequency channels at band edges showing periodic 'striping' behaviour (see Section 5.2.2.1). This conservative masking does not affect the overall noise profiles observed at any frequency sub-band, and reduces the pulse signal with respect to the noise at 1670.4 MHz.

Time-domain clipping significantly improves the prominence of the pulse profile at 1414.4 MHz, 1542.4 MHz and 1670.4 MHz. Overall noise levels are reduced for all three sub-bands. At 1542.4 MHz the pulse profile was only visible (at pulse phase ~ 0.2) after this processing stage. At 1670.4 MHz, clipping reduces the impact of correlated noise, flattening the off-pulse profile. No significant improvement is seen at 1350.4 MHz. Frequency-domain MAD flagging does not significantly affect profiles after clipping.

Figure 6.6 quantifies the impact of LOFT-e's RFI mitigation on each sub-band,
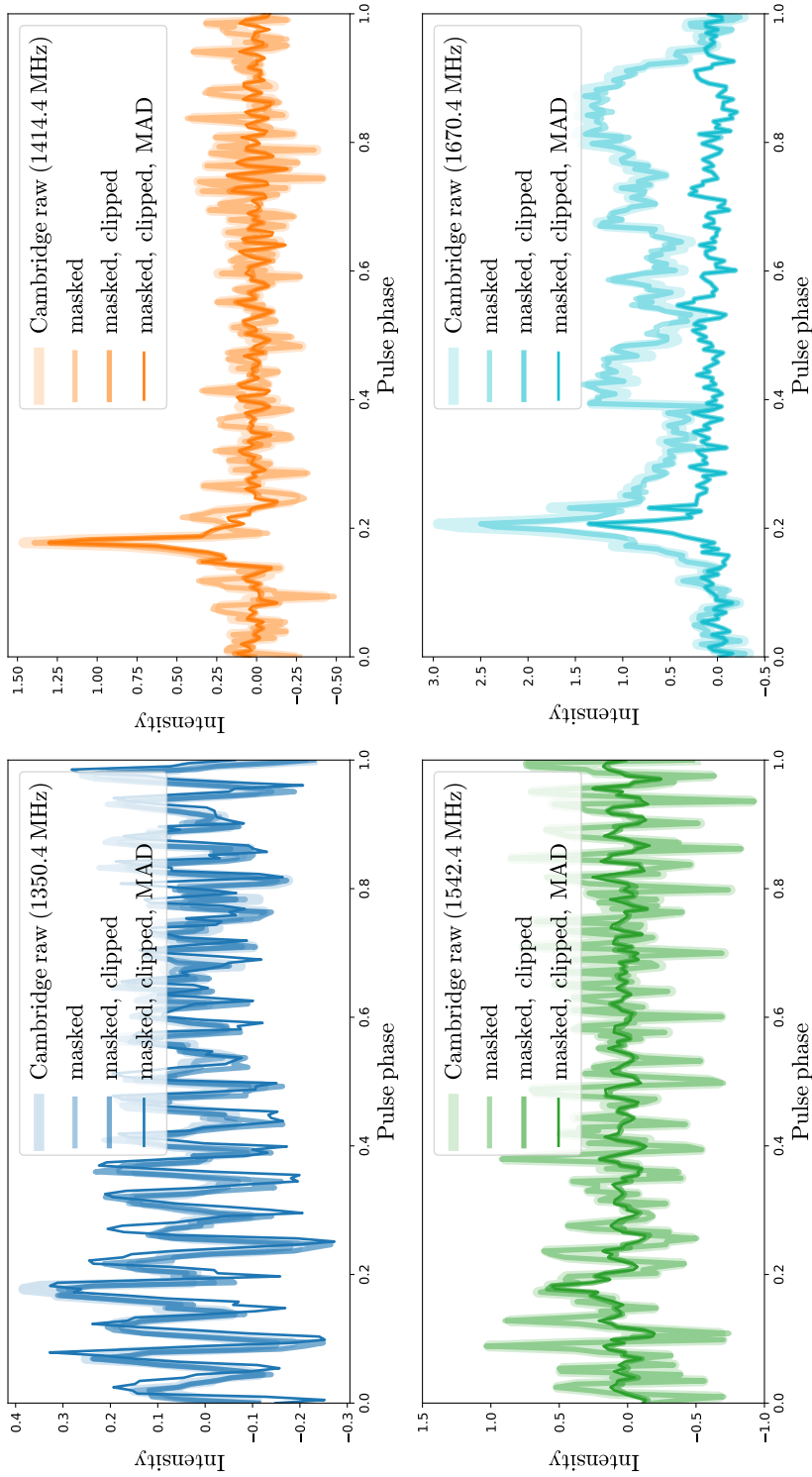
Figure 6.5: Intensity against pulse phase plots for PSR B0329+54, acquired using e-MERLIN's Cambridge element at different observing frequencies. Each pulse profile was formed from a LOFT-e filterbanked data product with 64 MHz bandwidth, 256 frequency channels, 0.25 MHz-spectral resolution and 64 μs temporal resolution. Sub-plots contain four lines which show the resulting pulse profiles formed after successive RFI-mitigation stages (see Chapter 5) are applied to data. Subplots for 1350.4 MHz, 1414.4 MHz, and 1542.4 MHz-centre frequencies are created from simultaneous 10-minute observations performed on 16/07/2018. The final subplot (1670.4 MHz) is created from a 1-minute observation performed on 18/01/2018.
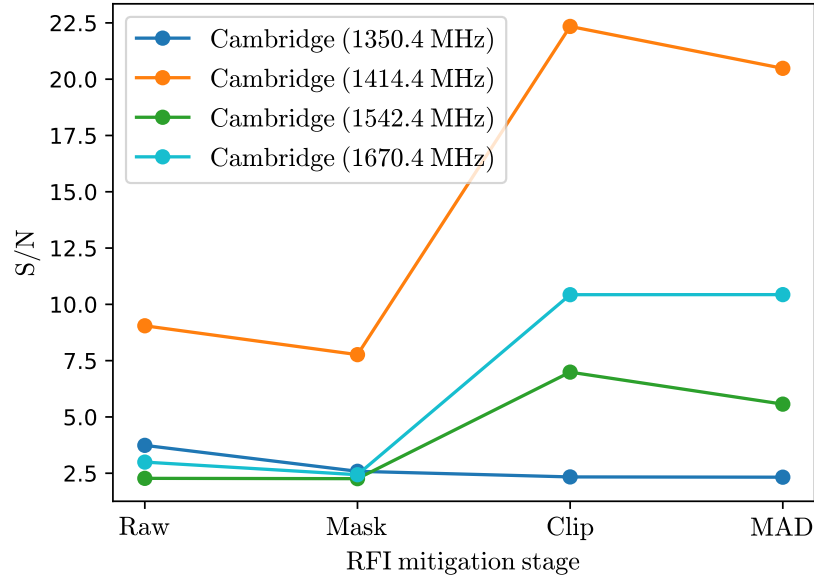
Figure 6.6: S/Ns recovered for the pulse profiles of PSR B0329+54 shown in Figure 6.6 after successive RFI mitigation stages (horizontal axis) have been applied to the data. Each line represents a different 64 MHz-wide sub-band centre frequency. Lines are visual aids and do not represent interpolation between individual points.

showing recovered peak S/Ns[2] for each sub-band after successive stages have been implemented. Masking slightly lowers S/N in all cases. Discarding the 1350.4 MHz-centred sub-band, which has no discernible pulse profile, time domain clipping consistently improves S/Ns by a factor > 2. Frequency domain MAD flagging has a minimal to negative impact on S/Ns.

These results indicate that e-MERLIN RFI has an inconsistent effect on performance at different sub-band centre frequencies. Of the three observations performed on 16/07/2018, 1414.4 MHz is the least contaminated band[3]. The minimal impact of MAD flagging also suggests that the majority of the RFI afflicting observations is in the time domain. During future observations frequency domain MAD flagging using a lower threshold should be implemented[4].
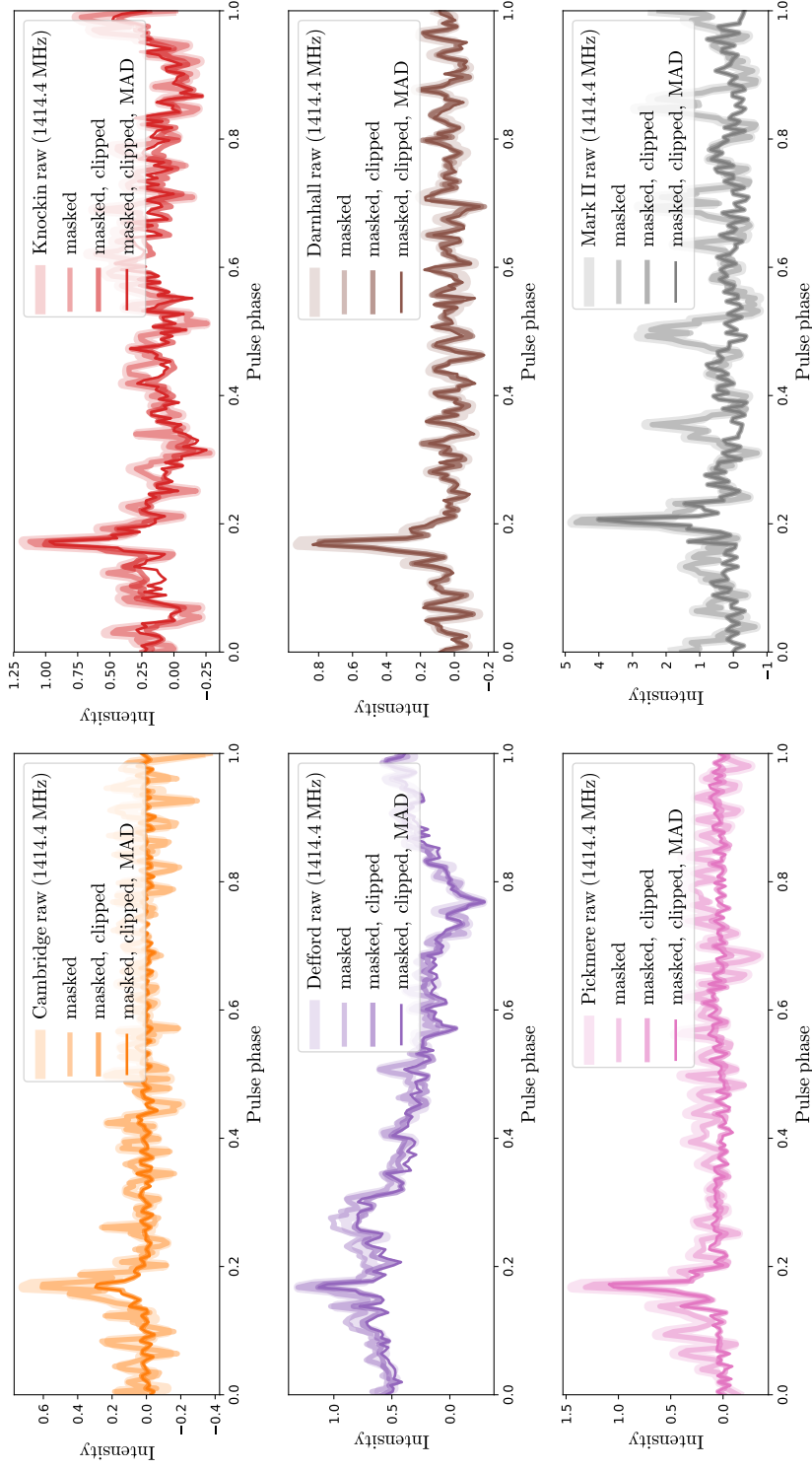
Figure 6.7: Intensity *vs.* pulse phase plots for PSR B0329+54, acquired using each e-MERLIN element for a 1414.4 MHz-centred frequency band. Cambridge, Knockin, Defford, Darnhall and Pickmere data are from a simultaneously-performed 10-minute observation on 16/07/2018. Mark II data are from a 1-minute observation performed on 18/01/2018. Figure style and data product details are as in Figure 6.5. Data was aquired at a different time to that in Figure 6.5.
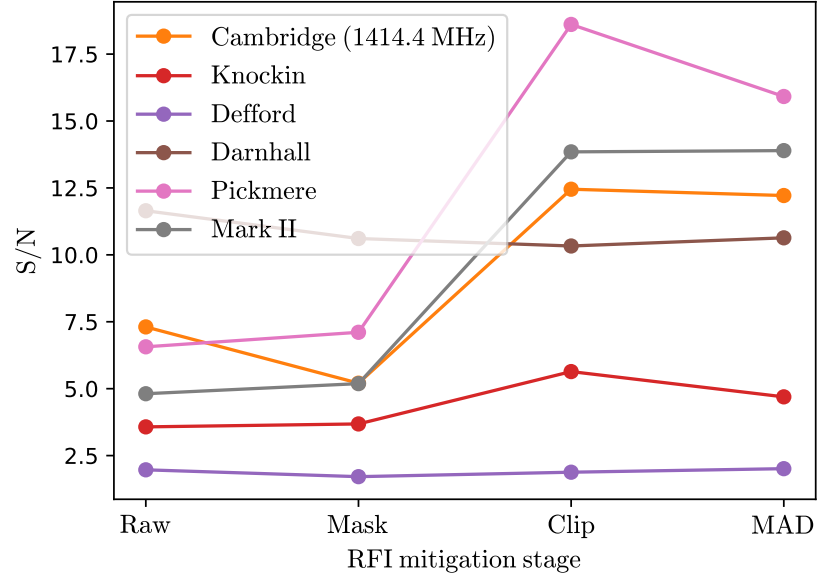
Figure 6.8: S/Ns recovered for the PSR B0329+54 pulse profiles shown in Figure 6.8 after successive RFI mitigation stages (horizontal axis) have been applied to data. Each line represents a different e-MERLIN telescope. Figure style after Figure 6.6.

6.1.2.2 PSR B0329+54 observations with multiple telescopes at 1414.4 MHz

Figure 6.7 presents intensity against pulse phase plots for B0329+54 acquired using all e-MERLIN telescopes at 1414.4 MHz. The Mark II profile is created using a 1-minute observation performed at the same time as the Cambridge (18/01/2018) observation in Section 6.1.2.1. All other profiles are created from simultaneous 10-minute observations performed 30 minutes after the Cambridge (16/07/2018) observations in Section 6.1.2.1.

Prior to RFI mitigation, noise levels vary significantly between telescopes. Cambridge, Darnhall and Pickmere show flat average off-pulse profiles, while Knockin, Defford and the Mark II indicate correlated noise. PSR B0329+54's pulse profile is visible in Cambridge, Knockin, Darnhall and Pickmere's data. In Defford and Mark II data the profile is partially obscured by correlated noise and would be difficult to identify without knowledge of the peak location provided by the other dishes.

---

[2]This peak S/N is the ratio of peak flux to noise, and may not represent the pulsar S/N.

[3]A direct comparison is not made between the 1350.4 MHz, 1414.4 MHz, and 1542.4 MHz-centred bands and the 1670.4 MHz-centre band as observations were of shorter duration and taken six months apart.

[4]Currently the MAD threshold is 3.

Masking does not offer significant improvements to pulse profiles. Time domain RFI clipping flattens off-pulse noise structure and reduces the amplitudes of noise fluctuations in Cambridge, Pickmere and Mark II data. Noise fluctuations are reduced for the Knockin and Defford profiles, however structure over the entire pulse phase remains. Frequency domain MAD flagging offers no significant profile improvements. Darnhall data remains consistently unchanged throughout the RFI mitigation process.

Figure 6.8 quantifies the impact of RFI mitigation on each telescope, showing peak S/N estimates after successive stages have been implemented. Masking improves S/Ns for Knockin, Pickmere and the Mark II data, and reduces S/Ns for Cambridge, and Darnhall data. Time domain RFI clipping improves S/Ns by a factor > 2 for Pickmere, Cambridge and the Mark II, improves Knockin's S/N more conservatively, and slightly reduces Darnhall's S/N. Application of MAD flagging has a minimal to negative impact on S/Ns as observed by downward trends.

These results indicate that on 16/07/2018, Pickmere's 1414.4 MHz-centre sub-band sometimes performs better post-masking and clipping than telescopes with lower SEFDs (e.g., Cambridge and Defford, see Table 4.1)[5]. This is most likely due to RFI contamination during the observation. Further evidence for time-dependent RFI contamination is the fluctuating S/N recovered for Cambridge's 1414.4 MHz-centred sub-band at different times of day (i.e. between Figures 6.6 and 6.8). PSR B0329+54 has a scintillation time (see Chapter 2) of ~ 500 s at 1500 MHz (Cordes & Lazio, 2002), so intensity fluctuations between these observations are expected. During both observations, raw data yields similar peak S/Ns (~ 9.0 and ~ 7.5 respectively). However, post-masking and clipping these diverge, increasing to ~ 22.5 and ~ 12.5 respectively. This indicates a varying RFI environment in the 30-minute period between observations which significantly affects the performance of the telescope. Together with the minimal effect on Defford and Darnhall data shown in Figure 6.8, these results suggest that current LOFT-e RFI mitigation techniques are not consistently effective over telescopes or time. We should note, however, that they generally only minimally reduce the detection significance of observable signals.

6.1.3 Demonstrating sub-band combination

Theoretically, an instrument's sensitivity to fast radio transients increases by a factor of the square root of its operational bandwidth $\Delta\nu$ (see Equation 3.7). LOFT-e's current

---

[5]The Mark II's sensitivity should not be directly compared to other telescopes here, as observations took place for different lengths of time and different days.
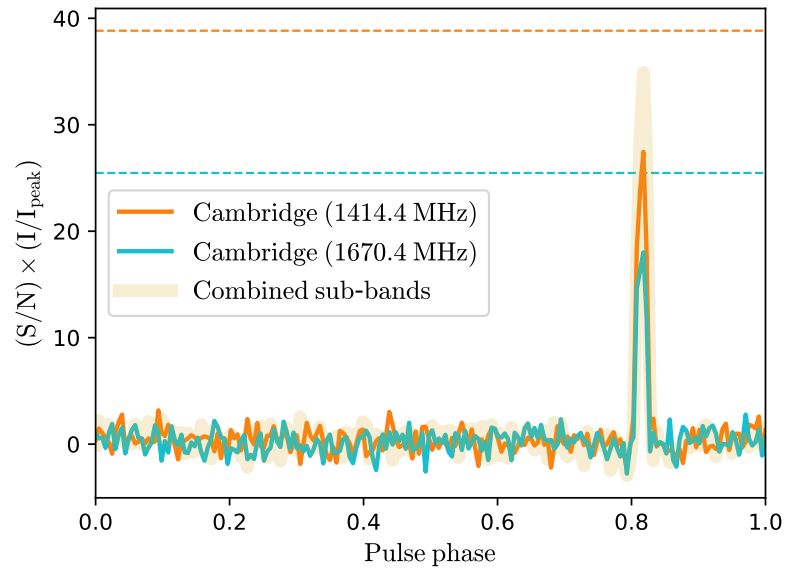
Figure 6.9: Folded pulse profile for PSR B1933+16 created from data acquired with e-MERLIN's Cambridge telescope on 02/11/2017. Average intensity normalised to I/RMS is plotted against pulse phase for visual clarity. Overplotted are profiles obtained from two 64 MHz-wide dynamic spectra of 10-minute duration with different centre frequencies (1414.4 MHz and 1670.4 MHz), and the profile obtained after combination of these sub-bands (solid lines). For reference, $\sqrt{2}$ sensitivity increases on each individual sub-band are indicated by dashed horizontal lines with colours corresponding to their profiles.
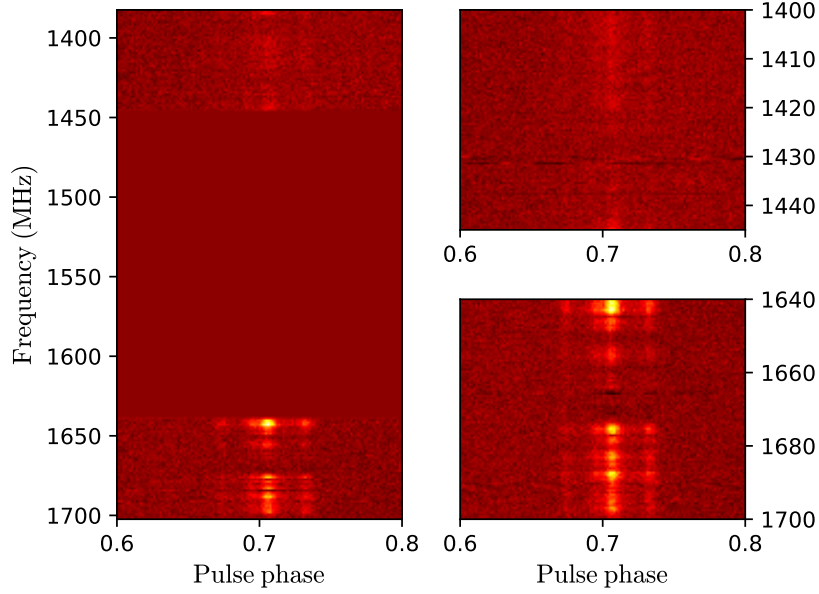
Figure 6.10: Folded frequency against pulse phase plots for a 10-minute observation of PSR B0329+54. The left subplot shows two combined sub-bands, observed simultaneously, centre on 1414.4 MHz and 1670.4 MHz. The right subplots show expanded views of these bands. Intensity fluctuations between 1660 MHz - 1670 MHz are consistent with scintillation effects.

optimum configuration enables the capture of 2×64 MHz sub-bands per telescope, affording a sensitivity increase of up to $\sqrt{2}$ over individual sub-bands. This is the maximum potential increase for combining sub-bands and depends on the data quality of the constituent components. Combination of a sub-band containing RFI contaminated non-Gaussian data (e.g. Cambridge's 1350.4 MHz-centre sub-band from Figure 6.5) with another sub-band will not result in a $\sqrt{2}$ sensitivity increase.

An example of the potential increase to sensitivity that may be obtained by combining e-MERLIN sub-bands is shown in Figure 6.9. This Figure presents normalised, folded pulse profiles for PSR B1933+16. The pulse profile is visible in both bands. Also included is the profile created from the summed combination of these sub-bands. The recovered combined profile's S/N falls between a $\sqrt{2}$ S/N increase on either individual band is consistent with the expected increase from averaging sub-bands which are not equally sensitive.

In addition to potential sensitivity improvements, increasing bandwidth allows for more complete study of spectral behaviours of any fast radio transients detected with LOFT-e. Figure 6.10 shows frequency against pulse phase plots for PSR B0329+54
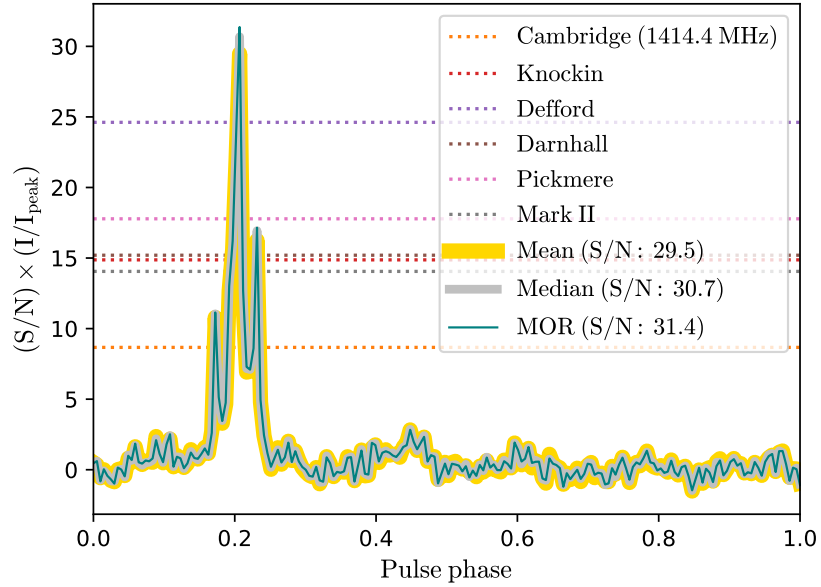
Figure 6.11: The folded pulse profile of PSR B0329+54 as observed on 18/01/2018. Average intensity (normalised to the pulse profile's peak S/N) is plotted against pulse phase for LOFT-e's three incoherent beamforming modes (solid lines; see Chapter 5). Beams were formed from 1-minute duration, 64 MHz-wide dynamic spectra centred on 1414.4 MHz for each of e-MERLIN's six dishes. Peak S/Ns obtained for the individual elements are indicated by dashed horizontal lines.

created after combination of two Cambridge sub-bands (centre on 1414.4 MHz and 1670.4 MHz). The intensity fluctuations visible between ~ 1660 to 1670 MHz as bright patches at pulse phase ~ 0.7 are consistent with the predicted NE2001 model scintillation bandwidth of PSR B0329+54 ($9.12^{+9.0}_{-3.5}$ MHz at 1500 MHz; Cordes & Lazio 2002).

6.1.4 Demonstrating incoherent beamforming techniques

Incoherent combination of e-MERLIN's constituent elements may afford an additional sensitivity increase of up to $\sqrt{6}$ (see Chapter 3) on a single dish, assuming identical dishes and RFI environments. When sensitivity and data quality differs between dishes this theoretical maximum will not be reached. The improvement in sensitivity afforded by LOFT-e's incoherent beamforming techniques is discussed in this section.

Figure 6.11 compares normalised, folded pulse profiles of PSR B0329+54 obtained by incoherent combination of e-MERLIN's six telescopes. The observation was
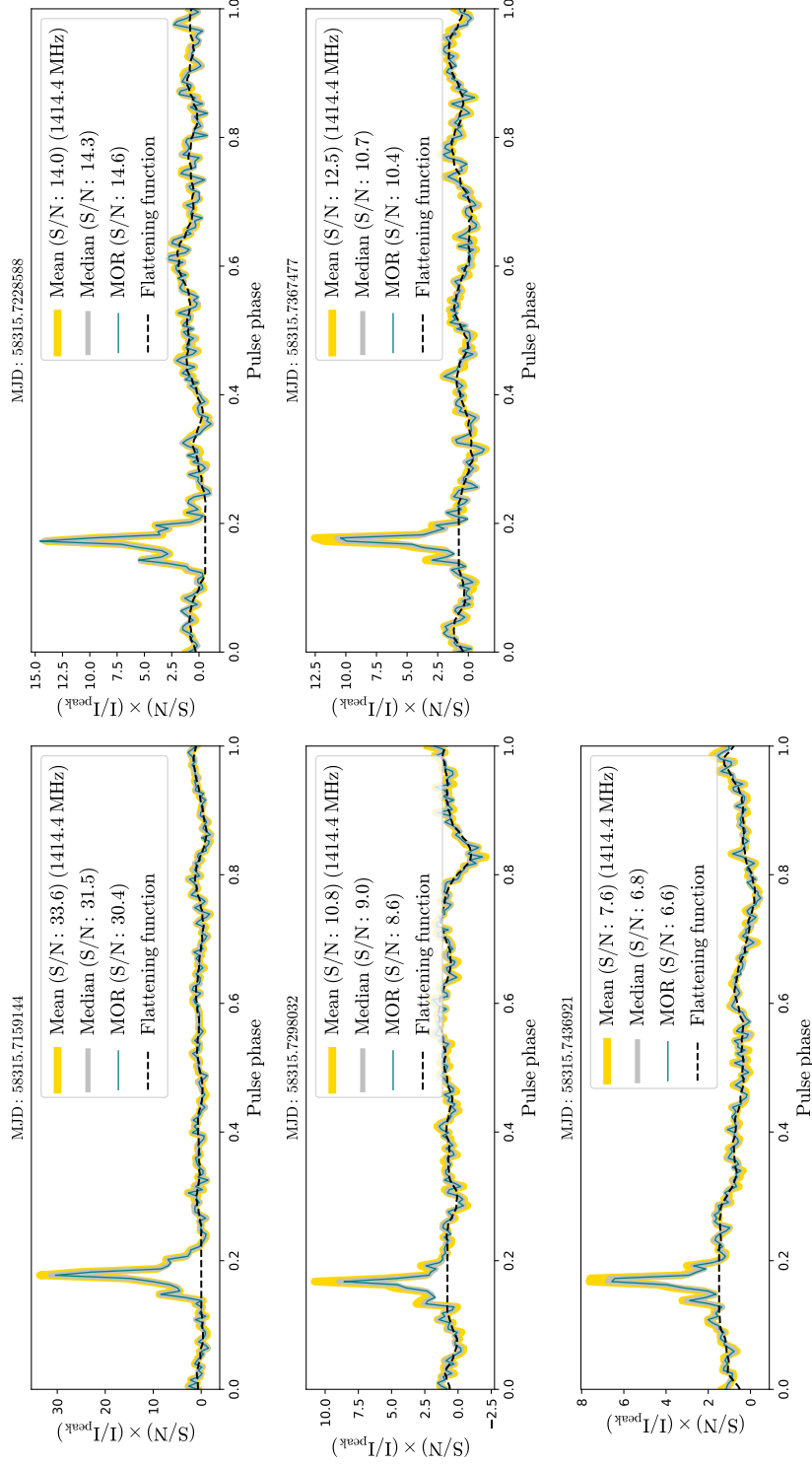
Figure 6.12: PSR B0329+54 pulse profile as observed over five consecutive 10-minute intervals on 16/07/2018. Average intensity (normalised to the pulse profile's peak S/N) is plotted against pulse phase for LOFT-e's three incoherent beamforming modes (solid lines; see Chapter 5). Beams were formed from five e-MERLIN elements (excluding the Mark II) on 1414.4 MHz. Beams were formed from 10-minute, 64 MHz-wide dynamic spectra centre on 1414.4 MHz. Dotted lines show a smoothed noise function which is subtracted from the pulse profile to obtain the more accurate S/Ns shown in Figure 6.13.
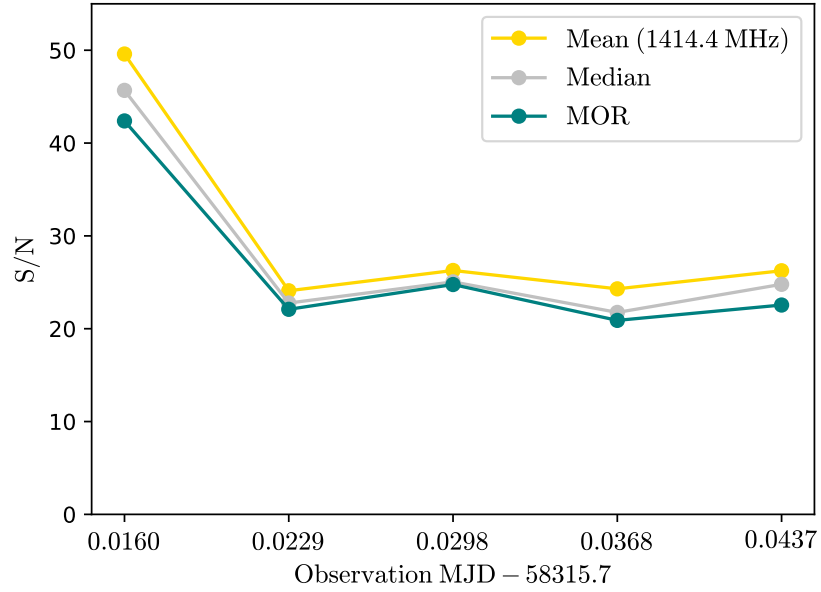
Figure 6.13: Peak S/Ns recovered for PSR B0329+54's five consecutive 10-minute observations in Figure 6.12 against time. The MJD on the time axis is offset by a constant value (58315.7) for visual clarity. S/Ns are recovered from pulse profiles after subtraction of the smoothed noise function inicated in Figure 6.12. Figure style after Figure 6.6.

1-minute in duration and performed on 18/01/2018. Masking and clipping were applied prior to beamforming. Each solid line indicates a beamforming mode described in Section 5.2.3. To demonstrate the improvement in S/N afforded by incoherent beamforming, the figure includes the S/Ns recovered for e-MERLIN's individual dishes as horizontal dashed lines. Figure 6.12 shows five similar curves formed from consecutive 10-minute datasets obtained on 16/07/2018[6].

For all observations, each beamforming technique applied yielded similar S/Ns. The final observation from 16/07/2018 (Figure 6.12, final subplot) is composed from data shown in Figure 6.7[7], so recovered S/Ns may be directly compared. The recovered standard incoherent beam sensitivity is only ~ 7.6 due to correlated noise contributed by Defford; a revised S/N calculated after subtraction of this trend (see Figure 6.13) is ~ 26.2. This marks a factor of $\sim \sqrt{2}$ improvement on the most sensitive individual dish (Pickmere). Likewise, the S/N improvement for the 18/01/2018 observation is a factor $\sim \sqrt{1.5}$ improvement on its most sensitive constituent telescope.

PSR B0329+54 has a frequency-dependent, multiple-component pulse profile, and

---

[6]In 16/07/2018 observations the Mark II was off-source and is excluded.

[7]Excluding the Mark II

exhibits mode-switching behaviour (Bartel et al., 1982). Despite sub-optimal sensitivity increases, post-RFI-mitigation and incoherent beamforming, noise in LOFT-e data is reduced to levels where multiple components are distinguishable. A prominent trailing component visible on 18/01/2018 but unseen on 16/07/2018 (see the difference between pulse profile shapes in Figures 6.11 and 6.12) indicates a mode change between the two observing sessions.

6.2   Single pulse detection with LOFT-e

By applying a $\sqrt{2}$ increase in sensitivity for incoherent beamforming, an expected FRB event rate at L-band of ~ 0.0013 FRBs day^{-1} is calculated for 1 ms, S/N $= 10$ events (see Equation 4.2). Applying anticoincidence techniques in order to spot simultaneous events in individual dishes may allow lower S/N FRBs to be identified and increase LOFT-e detection rates. However the expected event rate for S/N $= 5$ events is ~ 0.0033 FRBs day^{-1}, or < 1 event per 1700-hour e-MERLIN observing year. Improved sensitivity of e-MERLIN's incoherent beams is necessary to increase these rates. Investigation into methods which may minimise the impact of e-MERLIN RFI during observations are discussed in Section 6.3.1.

Localising sources of known Galactic single pulses (e.g., pulsars and RRATs) is another motivation for the development of LOFT-e. As discussed in Chapter 1, precise timing parameters are necessary to place RRATs on the P-$\dot{\text{P}}$ diagram (see Figure 1.2), and their localisation is required to successfully recover enough bursts for their accurate timing. Pulsars are commonly localised using gridding techniques (see Section 3.1.3), but due to the sporadic bursting nature of RRATs, observation times necessary to localise a single source using this technique may be unfeasibly large. The ability to localise RRATs with relatively few bursts may save on observing time and enable timing and high energy follow-up measurements (e.g., X-ray observations Zavlin & Pavlov 2004). This in turn will provide data required to solve the neutron star birthrate problem (which was introduced in Section 1.2.3).

The first RRAT single pulse detections by LOFT-e are presented in this section. Using time secured in recent e-MERLIN observing cycles, future observations of RRATs will allow detection of more bursts and demonstrate LOFT-e's ability to localise them via imaging. The method which will be used to do so is described at the end of this section.

6.2.1 LOFT-e detection of RRAT J1819−1458

Source	Purpose	Total time on source
J1331+305	Flux density calibrator	<10 minutes
J1935+1616	Target pulsar	31 minutes
J1407+284	Bandpass calibrator	<10 minutes
J1331+305	Flux density calibrator	28 minutes
J1819−1458	Target RRAT	4.8 hours
J1854−1539	Phase calibrator	55 minutes

Table 6.2: Sources observed during a LOFT-e RRAT single pulse detection experiment on 01/11/2017. Purposes for, and total time spent observing each source are provided.

On 01/11/2017, a proof-of-concept observation to demonstrate LOFT-e's single pulse detection pipeline was performed. The observing target, RRAT J1819−1458, was chosen for its emission of bright pulses at 1400 MHz (~ 3.6 Jy) and its relatively high burst rate ($\sim 18\,\mathrm{hr}^{-1}$)[8]. Precise timing parameters are already well known for this source (see Table 6.1).

Hu et al. (2011) have previously observed J1819−1458 with the 25 m Nanshan telescope at a central observing frequency of 1541.25 MHz. During these observations 423 bursts were detected over a period of 260 hours, 266 of which had S/Ns > 6. Using their results Hu et al. (2011) fit a power law flux intensity distribution to J1819−1458 bursts:

$$N(F) \propto F^{-\alpha}, \tag{6.2}$$

for N bursts of fluence F with index $\alpha = 1.6 \pm 0.2$. It follows from this number distribution that the total number of bursts which will be emitted by the source in 260 hours above a fluence threshold F_{thresh} is:

$$C_{\mathrm{thresh}} = \int_{F_{\mathrm{thresh}}}^{\infty} N(F)\,\mathrm{d}F \propto \frac{F_{\mathrm{thresh}}^{1-\alpha}}{1-\alpha} = A\frac{F_{\mathrm{thresh}}^{1-\alpha}}{1-\alpha}, \tag{6.3}$$

where A is a constant. Using the radiometer equation (Equation 3.7) and the Nanshan telescope's parameters[9], a fluence threshold $F_{\mathrm{min,Nan}} = 0.0092$ Jy s is calculated for the

---

[8]RRAT parameters are taken from the RRATalog: `http://astro.phys.wvu.edu/rratalog/`, (Accessed: 25/09/2018).

[9]The Nanshan telescope has SEFD $\simeq 307$ Jy, digitisation factor $\beta = \sqrt{\pi/2}$, $n_{\mathrm{p}} = 2$ polarisation
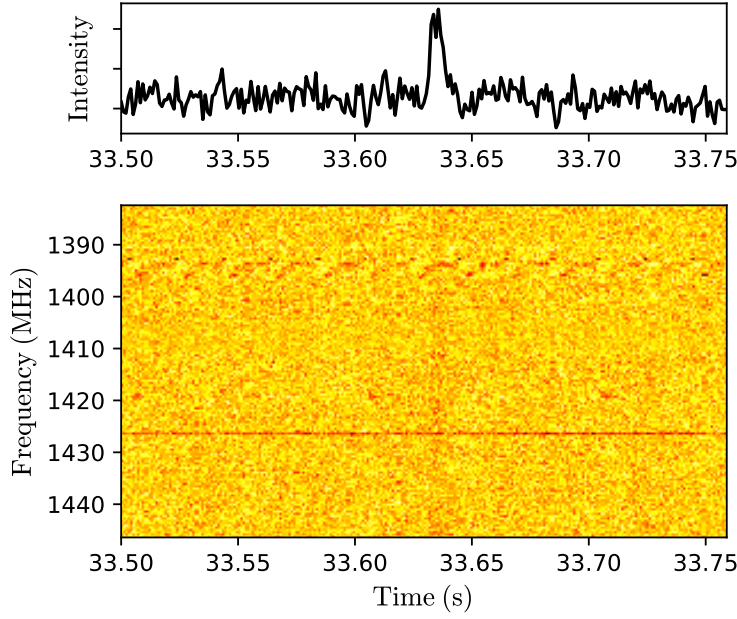
Figure 6.14: Dynamic spectrum of observing frequency against time for a single pulse candidate from RRAT J1819−1458 identified with LOFT-e (bottom) and its associated frequency-averaged timeseries (top). Data have been dedispersed to 196 pc cm^{-3}, and integrated to 1.024 ms temporal resolution for visual clarity. In the dynamic spectrum, colour indicates relative intensity. The burst, seen as a broad-band vertical line in the dynamic spectrum and a peak in the timeseries occurs at 33.64 s.

Nanshan telescope for $S/N = 15$, 4-ms-wide bursts. It therefore follows that:

$$A = 266 / \left(\frac{F_{\mathrm{min,Nan}}^{1-\alpha}}{1-\alpha} \right). \tag{6.4}$$

LOFT-e was configured to collect data from 3×64 MHz sub-bands during observations[10] for four telescopes[11]. Sub-bands chosen were previously identified to have relatively little RFI contamination. Due to correlator configuration issues, only data from Cambridge and Darnhall was recieved by LOFT-e. In total, 4.8 hours were spent on target. Using equations 6.3 and 6.4, and a 2-telescope e-MERLIN incoherent beam fluence threshold for $S/N = 8$[12], between 3 and 5 bursts were expected above LOFT-e's single pulse search threshold during the observation.

---

channels, an observing bandwidth $\Delta\nu = 320$ MHz, and a minimum detectable flux density $S_{\mathrm{min,Nan}} = 3.4$ Jy for $S/N = 5$ bursts at its sampling interval $\tau = 0.5$ ms (Hu et al., 2011).

[10]Sub-bands chosen were centred on 1414.4 MHz, 1606.4 MHz and 1670.4 MHz.

[11]The Mark II and Defford telescopes were unavailable.

[12]The current threshold for LOFT-e's single pulse search pipeline.

Collected data was processed using an early version of LOFT-e pipelines[13] Analysis of J1935+1616 data showed persistent noisy structure which hindered detection of the pulsar at 1606.4 MHz and 1670.4 MHz, therefore dynamic spectra created for RRAT observations were searched for single pulses on a sub-band by sub-band basis with an AstroAccelerate search threshold S/N = 8, below which candidates were discarded. Each dynamic spectrum was of 64 MHz bandwidth, 5-minute duration, with 64 μs temporal resolution and 0.25 MHz spectral resolution.

Figure 6.14 presents the dedispersed dynamic spectrum and associated frequency-averaged timeseries for a J1819−1458 burst candidate of S/N = 15. The candidate, found 33.64 s into the incoherent beam formed from e-MERLIN's 1414.4 MHz-centred sub-band data recorded three hours into observations at MJD 58059.6348384, was the only potential RRAT burst identified. The candidate's width of ∼ 4 ms and best dispersion measure ∼ 196 pc cm^{-3} are consistent with a burst from RRAT J1819−1458. The candidate was later identified in individual Cambridge and Darnhall 1414.4 MHz-centred data. The candidate was not recovered in data from other frequency sub-bands.

By applying Darnhall and Cambridge's best and worst-case SEFDs to the radiometer equation for a two-telescope incoherent beam, a fluence threshold $F_{\text{thresh,e-MERLIN}} = 0.027 \pm 0.012$ Jy s is calculated for 4-ms-wide, S/N = 15 bursts. Assuming equal FoVs for e-MERLIN's 2-dish incoherent beam and the Nanshan telescope, scaling the Nanshan burst rate ($R_{\text{Nan}} = 423/260$ hr^{-1}) to e-MERLIN's sensitivity (following Equation 4.2) yields an expected detection rate of 0.16 to 0.76 bursts of S/N = 15 per hour for e-MERLIN (i.e. 1 to 3 expected bursts during our observing period). Therefore the number of S/N=15 bursts emitted is consistent with our observations, however the total number of detected bursts was lower than expected. The varying emission rate of J1819-1458 (Bhattacharyya et al., 2018) and the impact of RFI on LOFT-e's incoherent beam sensitivity may contribute to this discrepancy.

In order to enable post-observation analysis of the quality of data recorded for each sub-band, the pulsar J1935+1616 was observed prior to this RRAT observation. To allow potential imaging of any bursts detected, observation time was spent alternating between the target RRAT and a phase calibrator source with a cadence of 7 minutes on target, 3 minutes on calibrator (slew times inclusive). A breakdown of time spent on each source during observations is shown in Table 6.2. However, due to a scheduling error the phase calibrator chosen for this observation was 8.4° from the target, while for

---

[13]In the version used, masking was performed using template masks on individual telescopes, however further RFI mitigation was applied post-incoherent beamforming. This saved on processing time as Python-based RFI clipping is slow.

phase coherence to be achieved, subject to atmospheric conditions, $< 5°$ is required[14].
Due to this error and lack of telescopes during the observation, no imaging of the burst
was attempted.

6.2.2   Future localising of RRAT bursts with LOFT-e

The detection of a burst from J1819−1458 demonstrates LOFT-e's ability to detect sin-
gle pulses from Galactic sources. To demonstrate LOFT-e's ability to localise RRATs,
a future observation of J1819−1458 will be performed using all e-MERLIN telescopes.
The times of arrival (ToAs) of pulses detected in LOFT-e's incoherent beams will be
used to extract appropriate visibilities from e-MERLIN correlated data products, which
will be used to image the source at radio wavelengths.

The e-MERLIN correlator is currently limited to 1 s integration times for visibil-
ities, which lowers its sensitivity to fast radio transients (see Chapter 3). Therefore
multiple bursts may need to be combined in order to image the source. The total num-
ber of bursts (and hence the total observing time) necessary to localise a source via
e-MERLIN imaging may be estimated. Following Equation 3.16, the peak S/N of a
radio pulse of flux density S_{peak} as measured by an N-element incoherent beam may
be written:

$$\mathrm{S/N_{peak,ib}} = \frac{S_{\mathrm{peak}}}{\mathrm{SEFD}} \times \sqrt{Nn_p\tau_{\mathrm{ib}}\Delta\nu_{\mathrm{ib}}}, \tag{6.5}$$

when the elements are identical with n_p polarisations, integration time τ_{ib}, and band-
width $\Delta\nu_{\mathrm{ib}}$. Similarly the S/N measured by a coherent image $\mathrm{S/N_{peak,ci}}$ is:

$$\mathrm{S/N_{peak,ci}} = \frac{S_{\mathrm{peak}}}{\mathrm{SEFD}} \times N\sqrt{n_p\tau_{\mathrm{ci}}\Delta\nu_{\mathrm{ci}}}, \tag{6.6}$$

for integration time τ_{ci}, and bandwidth $\Delta\nu_{\mathrm{ci}}$.

For a given instrument (e.g., e-MERLIN) illuminated by a particular burst, taking
the ratio of these quantities yields the signal to noise of the pulse in the coherent image
as a function of its signal to noise in the incoherent beam:

$$\mathrm{S/N_{peak,ci}} = \mathrm{S/N_{peak,ib}}\sqrt{N}\frac{\sqrt{\Delta\nu_{\mathrm{ib}}\tau_{\mathrm{ib}}}}{\sqrt{\Delta\nu_{\mathrm{ci}}\tau_{ci}}}, \tag{6.7}$$

It then follows from Equation 3.16 that the combination of N_b identical pulses results
in a signal with a S/N which scales as $\sqrt{N_b}$. Therefore when combining images from

---

[14]During future observations the phase calibrator J1825−1784, 2.76° from the target, will be used.

multiple identical bursts detected in an incoherent beam, the resulting S/N of the signal in the image is:

$$S/N_{peak,\sum ci} = S/N_{peak,ib} \sqrt{NN_b} \frac{\sqrt{\Delta v_{ib}\tau_{ib}}}{\sqrt{\Delta v_{ci}\tau_{ci}}}, \tag{6.8}$$

which may be rearranged to provide the number of identical bursts of a given S/N found in an instrument's incoherent beam which must be combined for the signal in the instrument's resulting image to reach a desired S/N:

$$N_b = \frac{1}{N} \left(\frac{S/N_{peak,\sum ci}}{S/N_{peak,ib}} \right)^2 \frac{\Delta v_{ci}\tau_{ci}}{\Delta v_{ib}\tau_{ib}}. \tag{6.9}$$

The e-MERLIN correlator uses its entire 512 MHz bandwidth during imaging ($\Delta v_{ci} = 512$ MHz). Currently, the shortest integration times offered by the correlator are of one second ($\tau_{ci} = 1$ s). Applying these and LOFT-e's ideal configuration of $N = 6$ telescopes, 128 MHz bandwidth and letting $\tau_{ib} = 4$ ms, approximately 7 bursts of $S/N_{peak,ib} = 15$ would need to be detected in LOFT-e's incoherent beams to form a 3σ signal in the image. The associated necessary time on-source is calculated to be between 4 and 5 hours.

This calculation assumes a consistent burst emission rate for J1819−1458, an RFI-free environment and equally sensitive e-MERLIN dishes. None of these assumptions are realistic, therefore extra time on source will be necessary to detect the appropriate number of bursts. Increased sensitivity and burst detection efficiency, or decreasing the number of bursts necessary for accurate localisation can improve LOFT-e's localisation capability. This may be achieved by identifying and removing sources of e-MERLIN's RFI (e.g. the periodic signals shown in Chapter 5), using software correlation to decrease imaging integration times (e.g., Law et al. 2011), or by using coherent beam-forming techniques to form multiple tied-array beams and grid individual bursts (e.g., Obrocka et al. 2015)

6.3 RFI investigation with LOFT-e

The existence of non-Gaussian RFI in data leads to decreases in telescope sensitivity. In order to mitigate RFI before it is sent to LOFT-e, e-MERLIN's AGC (introduced in Section 4.2.4) can be applied during requantisation of voltage data from 3.5-bit to 2-bit as part of the correlator's VDIF export mode. As discussed in Section 6.1.1, if

applied too frequently, periodic adjustment of the dynamic range of data may lead to issues recovering signals from pulsars or fast radio transients. Therefore the mode of application of the AGC was changed from continuous rescaling to a trigger-based system, which is applied after a source change during observations.

The improvements afforded by this new mode of AGC application have allowed the detection of pulsars and single pulses. However periodic RFI is still seen to occur in data recovered by LOFT-e, as shown in Section 4.5, and post-AGC application, the distributions of 2-bit voltage data have been seen to deviate from Gaussian profiles by varying amounts for different telescopes, polarisations and frequency sub-bands. This will contribute to sensitivity losses experienced by individual dishes.

This section presents investigations into the data quality of e-MERLIN dishes recorded by LOFT-e. These investigations aim to identify non-Gaussian behaviours in data from e-MERLIN telescopes which may help pinpoint sources of e-MERLIN's sensitivity loss. As multiple AGC triggers are sometimes necessary for optimal requantisation of data, and as fluctuating levels of RFI may influence the necessary cadence of AGC triggering, this work has allowed the best AGC application method during observations to be investigated.

The results of these investigations have been shared with e-MERLIN staff during over the course of LOFT-e's development and e-MERLIN's AGC code has subsequently been updated several times. Where updates have taken place between particular experiments is noted as appropriate.

6.3.1   Investigating the effects of manual AGC triggering

Tests were conducted in 2018 in order to investigate the outcome of manual AGC triggering on data distributions for different e-MERLIN telescopes and sub-bands centre frequencies. During these tests the following methodology was applied:

1. Record a brief period of data with LOFT-e. Do not trigger the AGC.
2. Cease recording and trigger the AGC.
3. Wait for a brief period (~ 10 s).
4. Repeat step 1.
5. Repeat steps 2 through 5 (three times).
6. Examine the Gaussianity of captured voltage data distributions.

Tests took place using a twelve-baseline board, 4-telescope setup (Cambridge, Pickmere, Knockin and Mark II)[15]. The frequency sub-bands captured were centred on 1286.4 MHz, 1414.4 MHz and 1542.4 MHz. No pulsar was targeted during this observation.

It is noted that this experiment took place on an e-MERLIN maintenance day during which Darnhall and Defford were unavailable. Further investigation with e-MERLIN staff confirmed that the Mark II and Pickmere telescopes were also under maintenance during the observations[16].

Figures 6.15, 6.16 and 6.17 compare the distributions of voltage data recovered prior to AGC triggering with distributions recovered after each successive AGC trigger. Three Figures are presented, each for data streamed to a different LOFT-e node during the experiment. As the voltage data recorded by LOFT-e is 2-bit (see Chapter 4 for details on e-MERLIN's data format), each recorded sample may occupy one of four potential bins. In each plot, cross-hatched histograms provide a reference for an ideally Gaussian-distributed data set (where 68% of data occupies the inner two bins and the remaining data occupies the outer two bins).

After each successive AGC trigger, Cambridge data for all recorded sub-bands and both polarisations[17] (see Figure 6.15, top 3 rows) and Knockin polarisation 0 data (see Figure 6.16, top 3 rows) assume more Gaussian profiles. This is demonstrated by each approaching the ideal distribution shown by the cross-hatched area. Slight variations in distribution occur between individual triggers. Knockin's polarisation 1 data, however, does not vary significantly between triggers for any sub-band. The majority of its data are concentrated in the third bin. Pickmere (all Figures, bottom row) and the Mark II (Figure 6.17, top 3 rows) show no deviations between AGC triggers and all of their data for all recorded data sets and sub-bands is concentrated in the third bin.

As maintainence may have affected the Mark II and Pickmere dishes, their data distributions may have been affected prior to entering the correlator. Results for these two elements are inconclusive. Due to this and the inconsistencies in behaviour observed between Cambridge and Knockin, the experiment was scheduled for repetition on a non-maintenance day.

---

[15]Darnhall and Defford were unavailable due to maintenance.

[16]And that Pickmere data was mapped to a single bin while still in 3.5-bit format, prior to AGC scaling and VDIF export mode.

[17]Pol 0 distributions are indicated in filled blue, pol 1 distributions are indicated by orange outlines.

Figure 6.15: Normalised histograms of LOFT-e captured 2-bit voltage data distributions within their four potential bins. Plots show data acquired with e-MERLIN's Cambridge and Pickmere elements. Each row shows data for a telescope and sub-band centre frequency prior to manual AGC triggering (subplot 1) and after successive AGC triggers (subplots 2 through 4). Individual subplots show normalised counts against bin number. Each contains one second of data histogrammed for each polarisation (pol 0: filled blue, pol 1: unfilled yellow). A cross-hatch of an ideal Gaussian distribution is provided for reference. Running means and standard deviations of the data (averaged every 32,000 timesamples) are inset (upper left and upper right respectively).

Figure 6.16: Normalised histograms of LOFT-e captured 2-bit voltage data distributions within their four potential bins. Plots show data acquired with e-MERLIN's Knockin and Pickmere elements. Subplot style follows Figure 6.15.

Figure 6.17: Normalised histograms of LOFT-e captured 2-bit voltage data distributions within their four potential bins. Plots show data acquired with e-MERLIN's Mark II and Pickmere elements. Subplot style follows Figure 6.15.

6.3.1.1 Manual triggering with updated AGC code

Section 6.3.1's experiment was reprised on a non-maintenance day when all six e-MERLIN antennae were operational. Only eight functional baseline boards were available during this experiment, therefore an eight-baseline board, 6-telescope configuration was used. The combinations of frequency sub-bands which may be recorded simultaneously are dependent on which baseline boards are operational. Therefore the exact configuration used in Section 6.3.1 could not be replicated.

For all telescopes, the frequency sub-band centered on 1414 MHz was captured. The remaining two baseline boards streamed sub-bands centred on 1350.4 MHz and 1542.4 MHz for Cambridge. The observing target during this test was PSR B0329+54[18]. The method followed was identical to that in Section 6.3.1, however AGC triggering was performed five times[19].

Figures 6.18, 6.19 and 6.20 compare the distributions of VDIF data recovered prior to AGC triggering with distributions recovered after each successive AGC trigger. The layout of each sub-plot is identical to those in Section 6.3.1. Three plots are provided, each for data streamed to a different LOFT-e node during the experiment.

Prior to AGC triggering, data recorded for all telescopes, polarisations and sub-band centre frequencies, is 100% contained in a single bin. In all cases, data assumes a distribution closer to the ideal Gaussian profile upon a single AGC trigger. In all cases except for Cambridge sub-bands (Figure 6.20), distributions remain stable after each successive AGC trigger. Cambridge sub-bands take two AGC triggers to reach their most Gaussian-like profile, after which Cambridge's 1414.4 MHz and 1542.4 MHz-centred sub-bands remain stable. Cambridge's 1350.4 MHz-centred sub-band fluctuates noticeably around a near-Gaussian profile with each successive trigger.

In all cases except for Knockin and Defford's 1414.4 MHz-centred sub-bands (Figure 6.18, top two rows) and Cambridge's 1350.4 MHz-centred sub-band (Figure 6.20, bottom row), data distributions appear consistent between polarisations. Knockin and Defford's polarisation 1 data display distributions heavily weighted to the third bin after all AGC triggers. Cambridge (1350.4 MHz) polarisation 1 data fluctuates wildly between all data being concentrated in the outer two or inner two bins.

---

[18]During these observations the Mark II telescope was under e-MERLIN's Pulsar Control Mode and therefore was not on source. This has no impact on AGC test results, which are only concerned with the profiles of data.

[19]Following previous LOFT-e results and further investigation by e-MERLIN staff, e-MERLIN's AGC code was found to behave inconsistenly between telescopes, sub-bands and baseline boards. Between the time of this and the previous experiment, e-MERLIN's AGC code was upgraded to a newer version supplied by NRAO and installed by e-MERLIN staff.

Figure 6.18: Normalised histograms of LOFT-e captured 2-bit voltage data distributions within their four potential bins. Plots show data acquired with e-MERLIN's Knockin, Defford and Darnhall elements. Each row shows data for a telescope and sub-band centre frequency prior to manual AGC triggering (subplot 1) and after successive AGC triggers (subplots 2 through 4). Subplot style follows Figure 6.15.
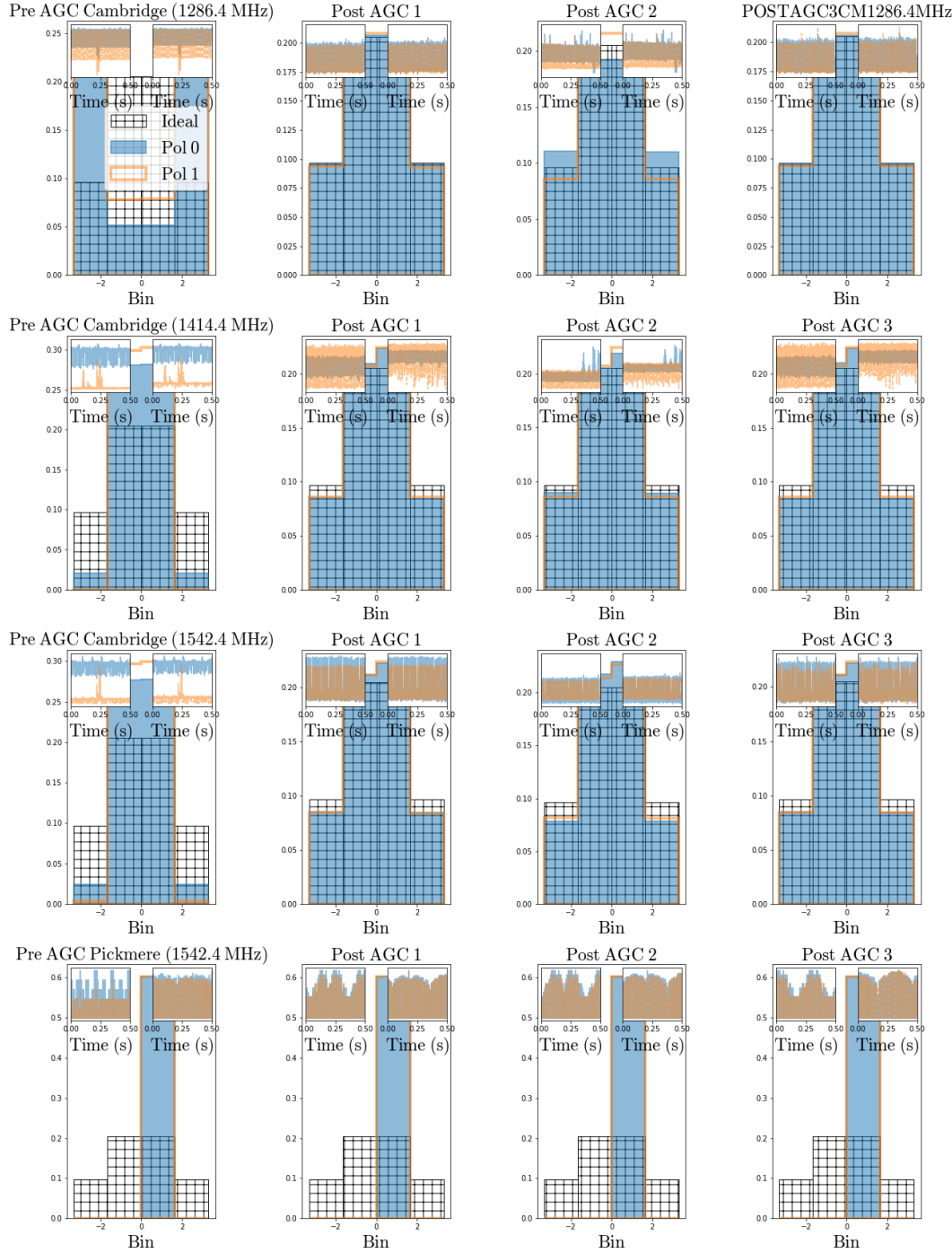
Figure 6.19: Normalised histograms of LOFT-e captured 2-bit voltage data distributions within their four potential bins. Plots show data acquired with e-MERLIN's Pickmere and Mark II elements. Each row shows data for a telescope and sub-band centre frequency prior to manual AGC triggering (subplot 1) and after successive AGC triggers (subplots 2 through 4). Subplot style follows Figure 6.15.
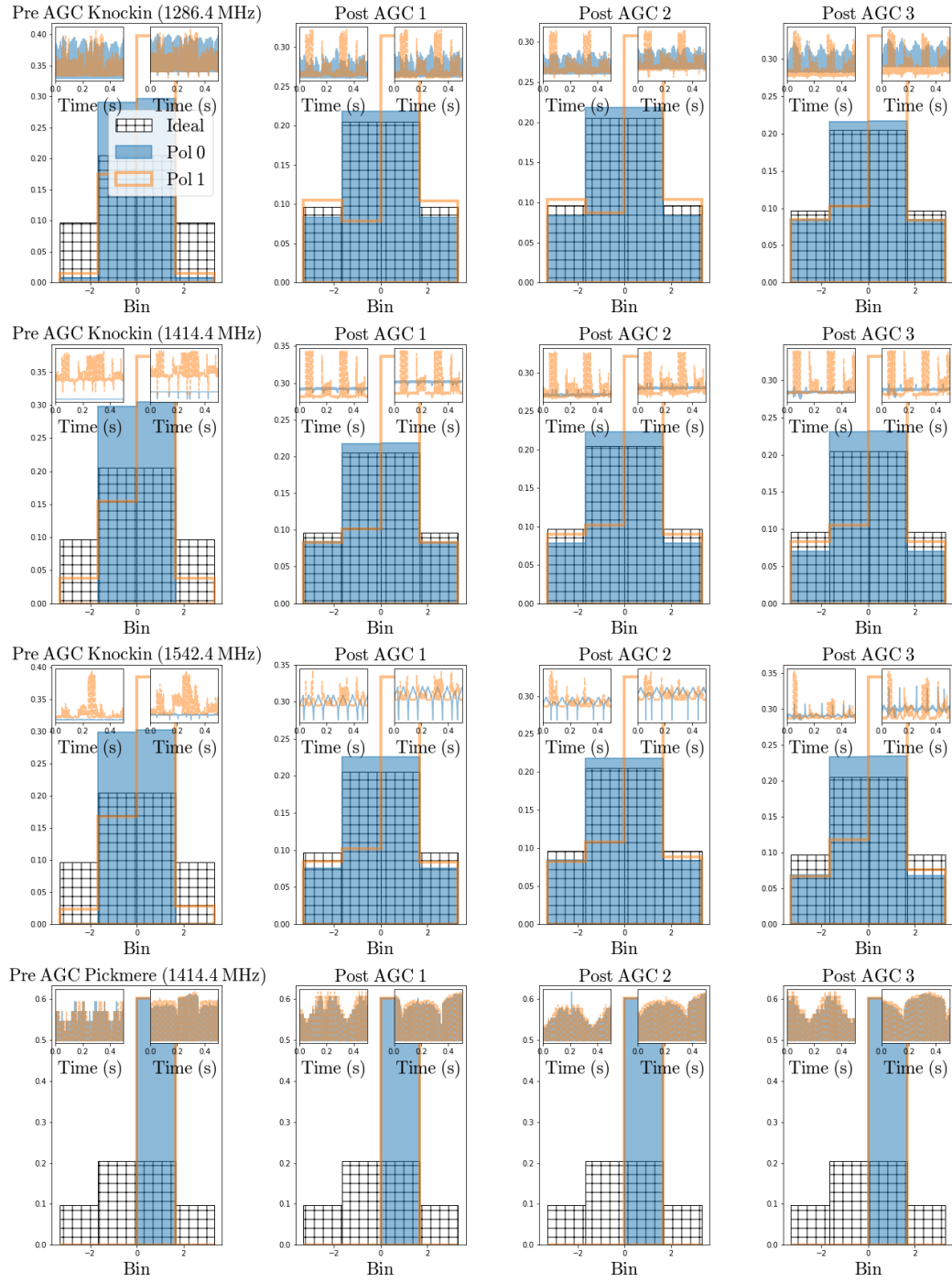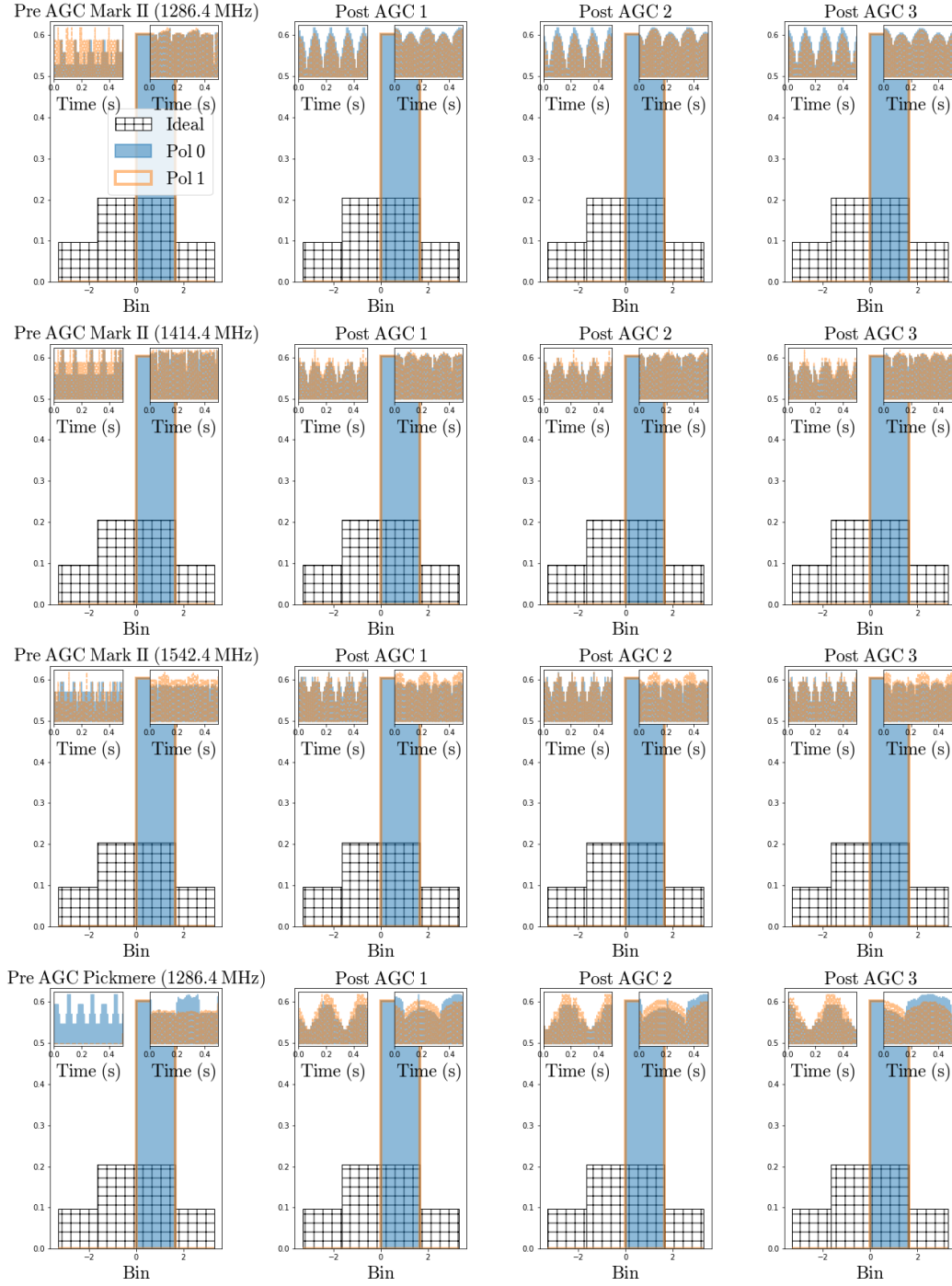
Figure 6.20: Normalised histograms of LOFT-e captured 2-bit voltage data distributions within their four potential bins. Plots show data acquired with e-MERLIN's Cambridge element. Each row shows data for a different sub-band centre frequency prior to manual AGC triggering (subplot 1) and after successive AGC triggers (subplots 2 through 4). Subplot style follows Figure 6.15.

Only Knockin's 1414.4 MHz and Cambridge's 1542.4 MHz sub-bands were recorded both for this experiment and Section 6.3.1's experiment. Plots which may be compared are Figure 6.16, row 2 with Figure 6.18, row 1 for Knockin data; and Figure 6.15, row 3 with Figure 6.20, row 2 for Cambridge data. After multiple AGC triggers, data distributions seem consistent between the two experiments. In Section 6.3.1 it did not take multiple triggers for Cambridge data to reach optimal distributions. Knockin data behaves consistently between every AGC trigger in both experiments.

Due to inconsistencies between experiments, baseline boards, telescopes and frequency sub-bands, it is difficult to determine whether any improvement was afforded by the upgraded AGC. As each of e-MERLIN's observing bands consist of eight sub-bands, a more complete analysis of the AGC's effect on e-MERLIN's data distributions is advisable. Investigation into the origin of consistently biased distributions of polarisation 1 data in Defford and Knockin's 1414.4 MHz sub-bands should also take place.

6.3.2   The stability of data distributions after AGC triggering

As shown in Section 6.3.1.1, multiple AGC triggers were necessary in order to attain Gaussian data distributions for Cambridge sub-bands (Figure 6.20), and AGC triggering did not always optimally redistribute data (see, e.g., variations in polarisation 1 behaviour between Cambridge's 1350.4 MHz sub-plots). Therefore automated AGC triggering may not always result in maximised sensitivity during observations. During future observations, manual triggering of the AGC until optimal distributions are achieved followed by no further AGC triggering may be a preferable approach. To explore this option, the stability of data distributions on a long timescale post-AGC triggering was investigated.

Following Section 6.3.1.1's experiment, no further AGC triggering was initiated. Data was captured at semi-regular intervals over a ~ 160 minute period until the end of the observation time window. The results are presented in Figures 6.21, 6.22 and 6.23, which compare the distributions of one second of data for each telescope and sub-band centre frequency at nine intervals post-final AGC triggering. Individual subplots are laid out identically to those in the previous two sections. Three plots are provided, each for data streamed to a different LOFT-e node during the experiment. As this experiment took place immediately after Section 6.3.1.1's tests, mapping between telescopes, sub-bands and LOFT-e nodes is identical.

Data distributions remain stable at each measured interval for all telescopes and sub-bands excluding for the Mark II's 1414.4 MHz-centred sub-band and Cambridge's
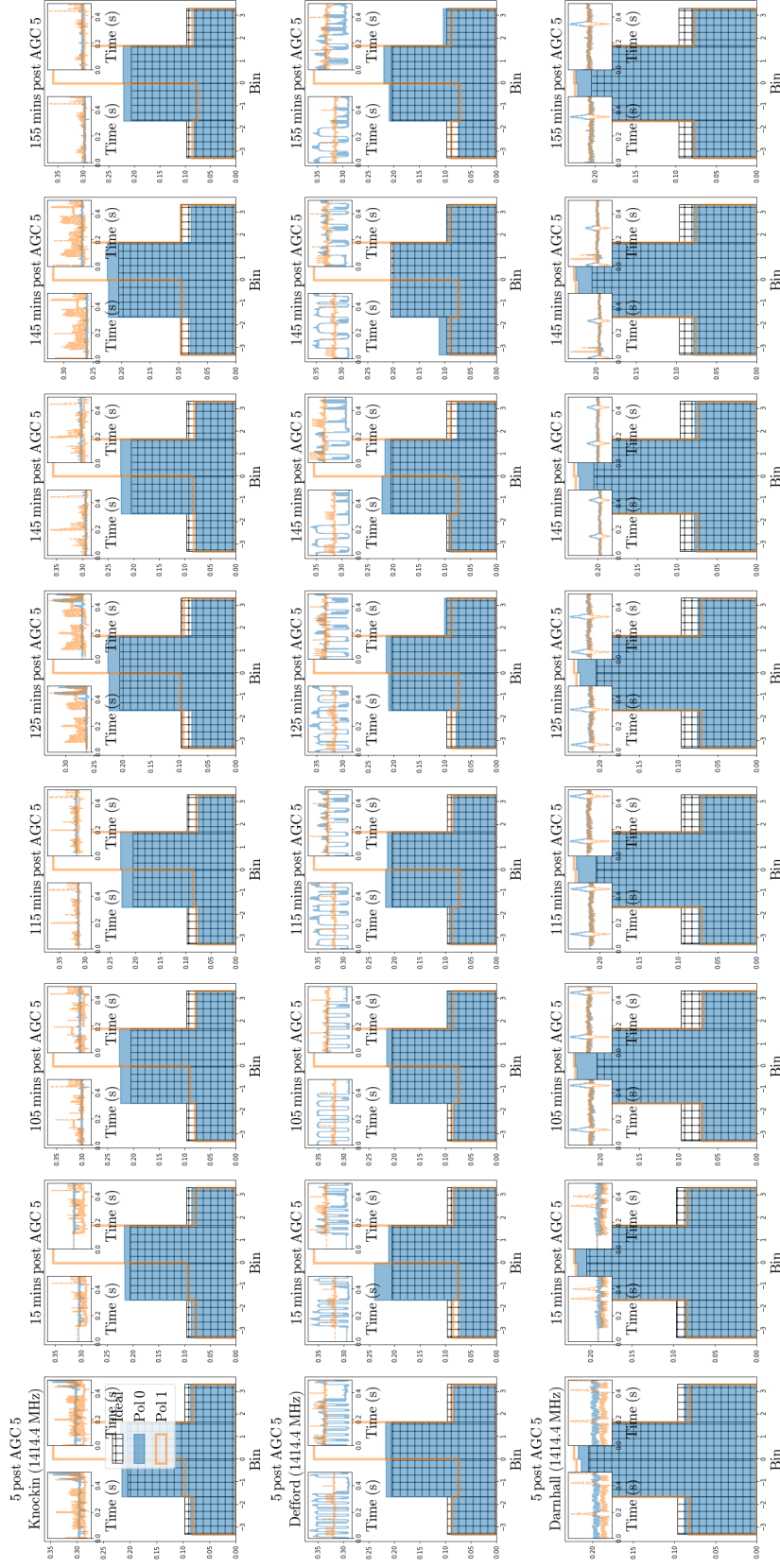
Figure 6.21: Normalised histograms of LOFT-e captured 2-bit voltage data distributions within their four potential bins. Plots show data acquired with e-MERLIN's Cambridge element. Each row shows data for a telescope and sub-band centre frequency. Columns show distributions measured at successive time intervals after the final AGC trigger detailed in Section 6.3.1 was performed. Subplot style follows Figure 6.15.
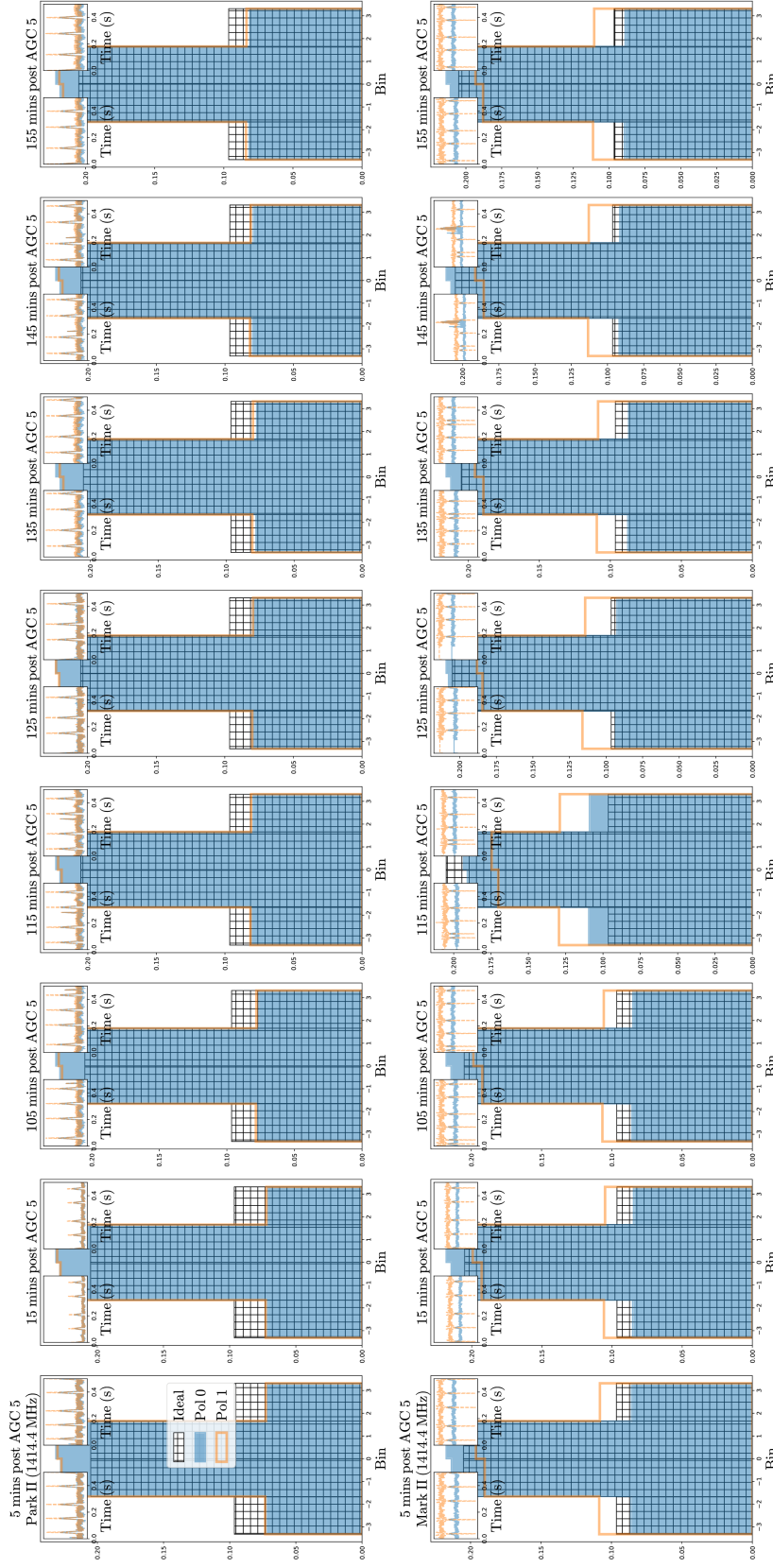
Figure 6.22: Normalised histograms of LOFT-e captured 2-bit voltage data distributions within their four potential bins. Plots show data acquired with e-MERLIN's Pickmere and Mark II elements. Each row shows data for a telescope and sub-band centre frequency. Columns show distributions measured at successive time intervals after the final AGC trigger detailed in Section 6.3.1 was performed. Subplot style follows Figure 6.15.
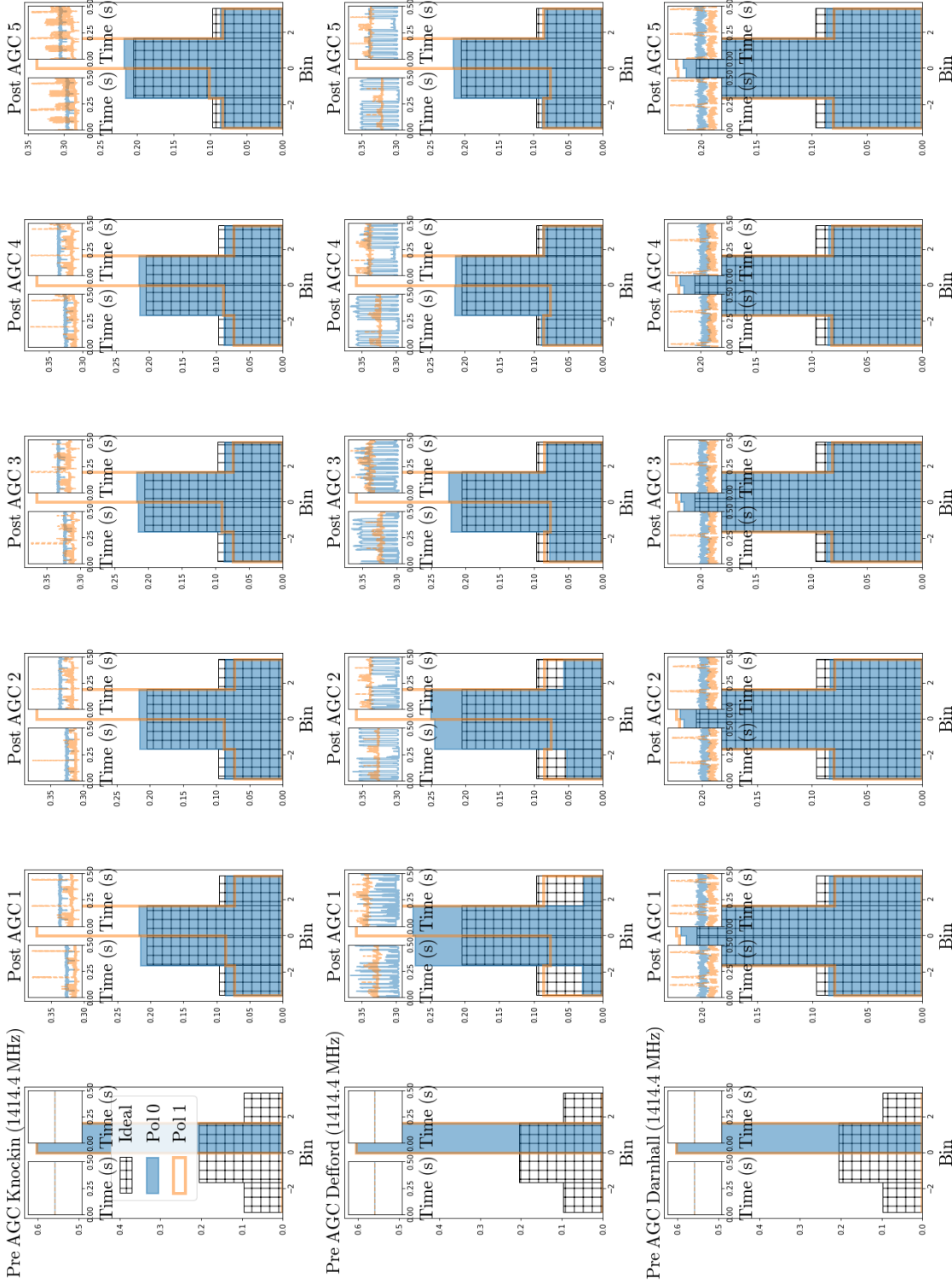
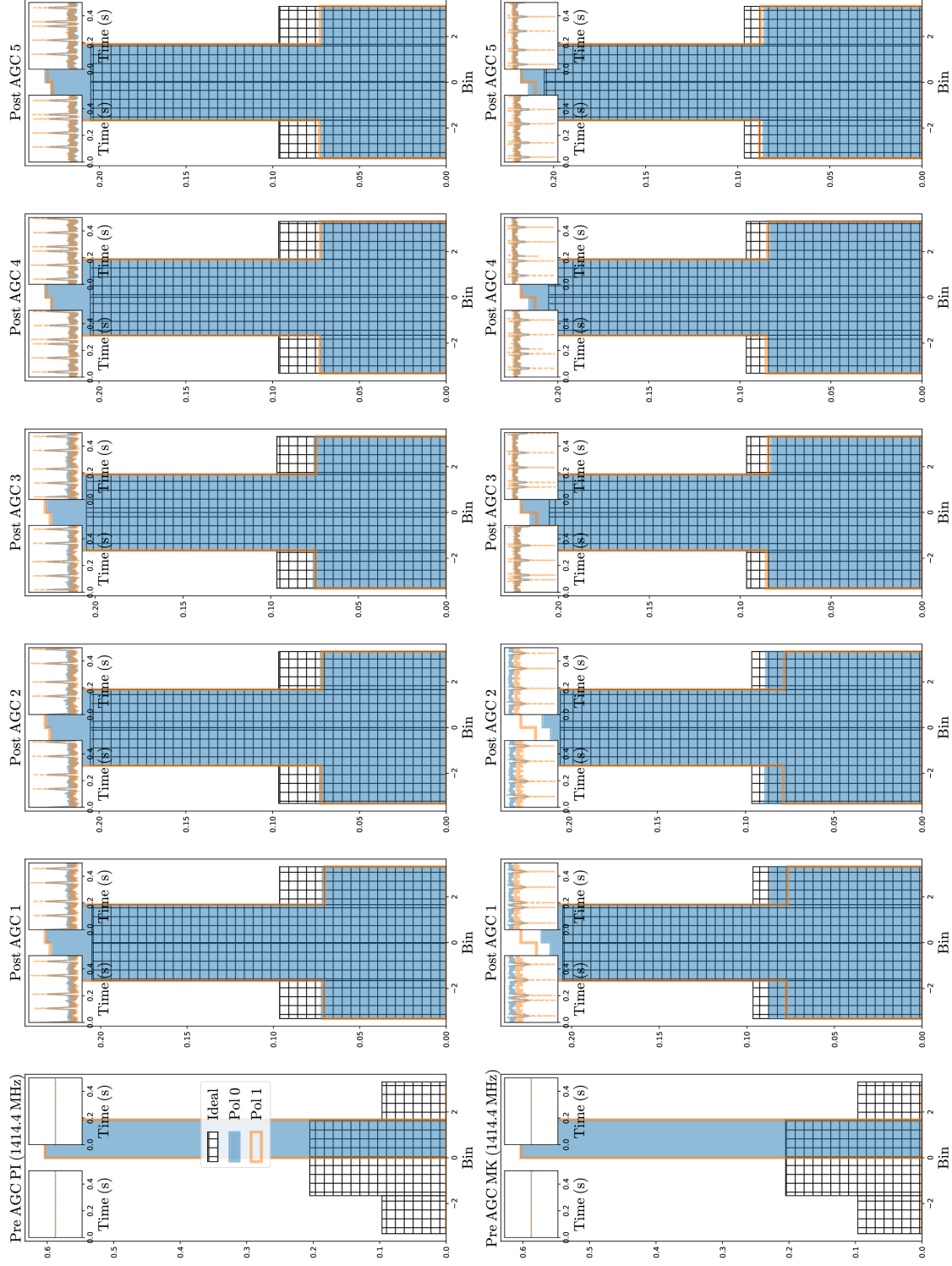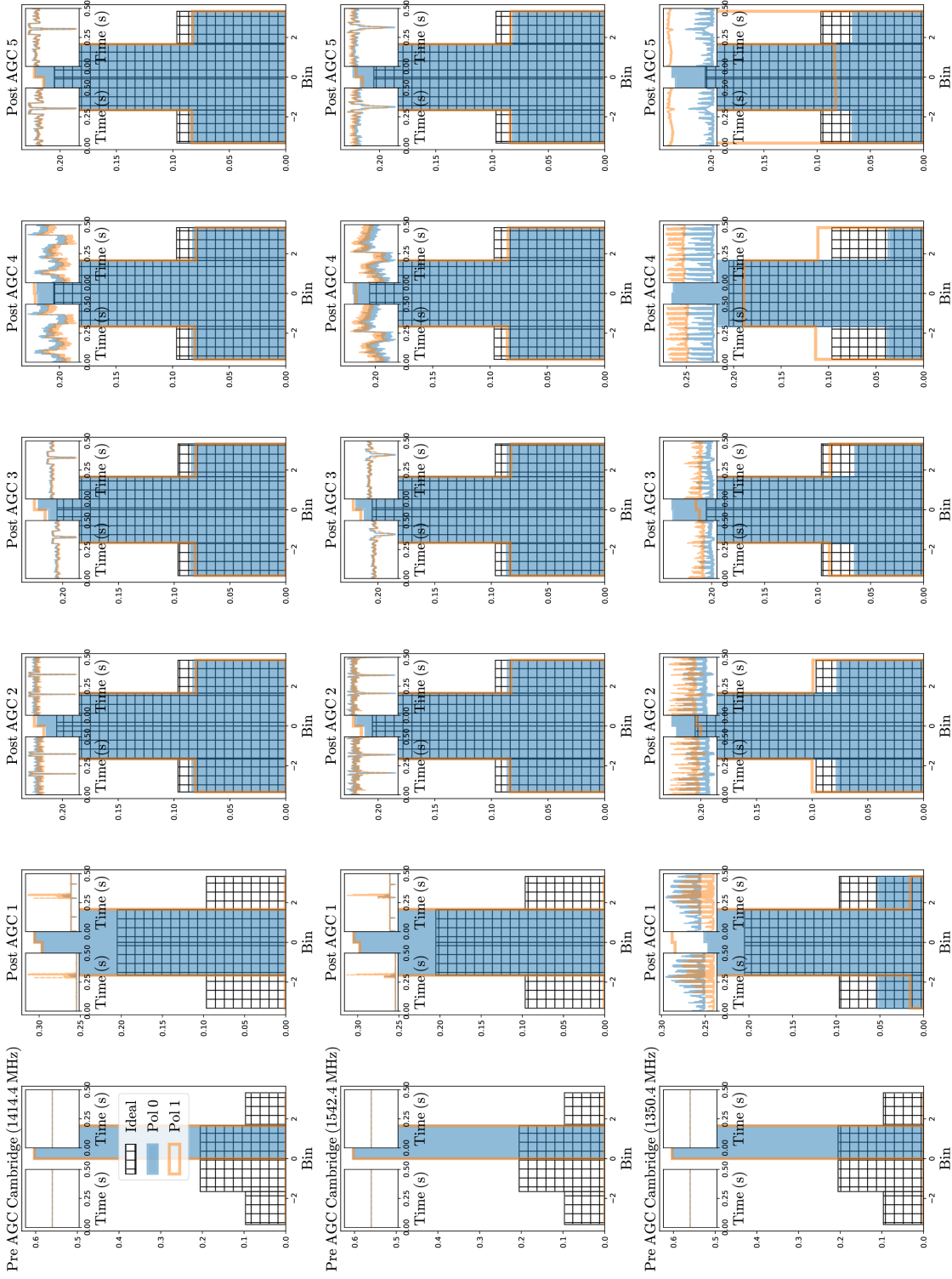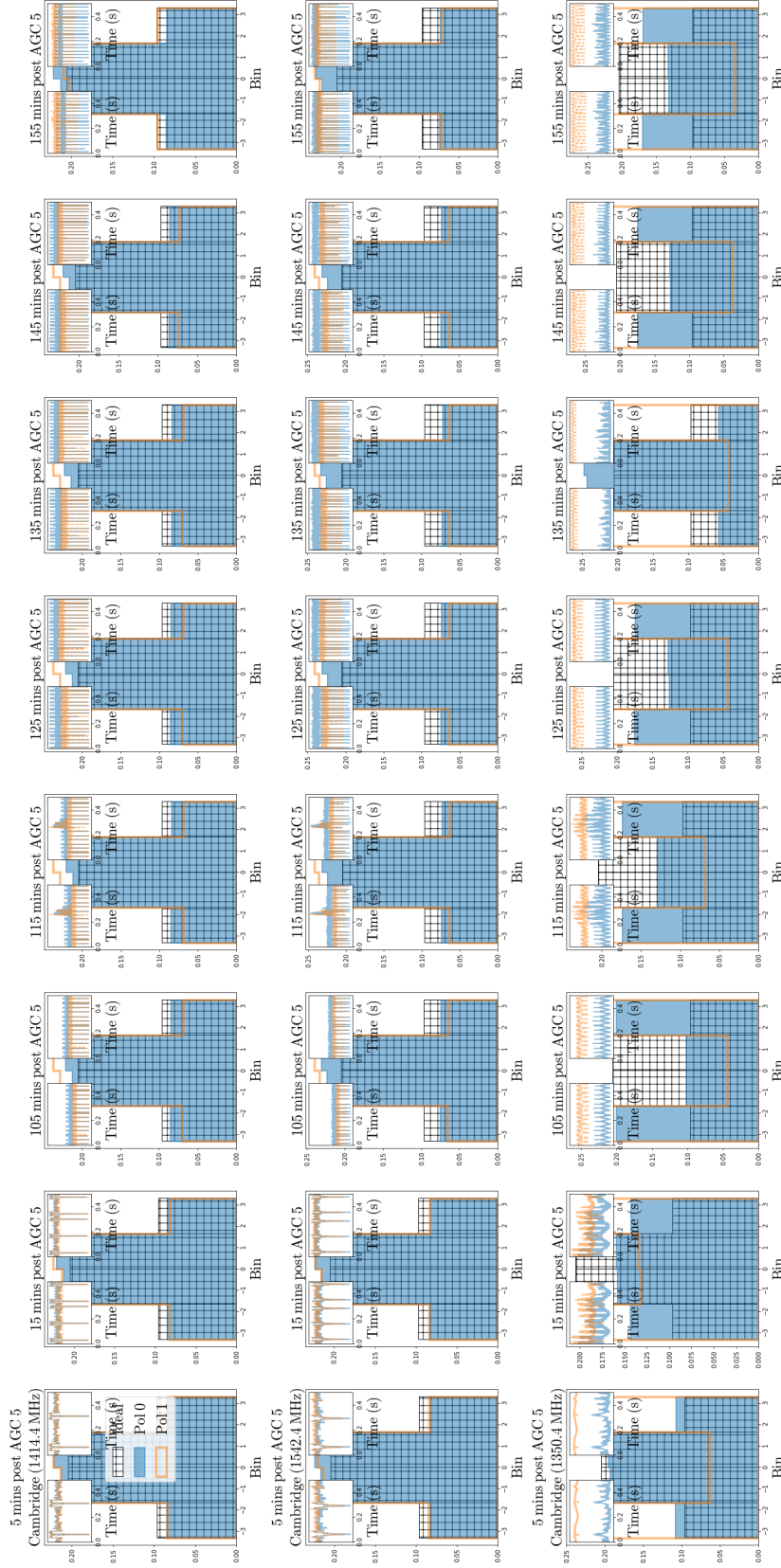**Figure 6.23:** Normalised histograms of LOFT-e captured 2-bit voltage data distributions within their four potential bins. Plots show data acquired with e-MERLIN's Knockin, Defford and Pickmere elements. Each row shows data for a telescope and sub-band centre frequency. Columns show distributions measured at successive time intervals after the final AGC trigger detailed in Section 6.3.1 was performed. Subplot style follows Figure 6.15.

1350.4 MHz-centred sub-band. The Mark II's polarisation 0 data (Figure 6.22, bottom row) fluctuates around the ideally distributed case, however distributions remain similar to that of an ideal Gaussian. Cambridge's 1350.4 MHz-centred sub-band (Figure 6.23, bottom row) however, fluctuates significantly from interval to interval for both polarisations. Its polarisation 0 data begins close to a Gaussian (see subplot 1). After 15 minutes (see subplot 2), data appears almost evenly distributed between bins, following which more data begins to occupy the outer two bins (see subplots 3 to 5). After 135 minutes (see subplot 6), data assumes a Gaussian profile, before deviating once more (see subplots 7 to 9). Polarisation 1 data consistently favours the outer two bins throughout the observation. As no AGC triggering occured between these observations, these results indicate a highly variable RFI environment for Cambridge at an observing frequency of 1350.4 MHz. This result is consistent with the lack of discernable pulse profile for PSR B0329+54 during the observation earlier on in the day at 1350.4 MHz (see Figure 6.5).

6.3.3 Implications for AGC triggering

For all data acquired by LOFT-e and presented in Section 6.1.1, pulse profiles associated with the observing target pulsars could not be recovered. Analysis of time-series data uncovered systematic non-Gaussian behaviour which was attributed to the e-MERLIN correlator by comparison of data aquired during a single observation by the Lovell Telescope which was acquired both prior to and post-correlator. By disabling e-MERLIN's AGC's continual data scaling and instead triggering the AGC prior to individual observations, the behaviour of data was improved which subsequently enabled the detection of pulsars (as shown in Section 6.1.2).

The impact of the AGC on LOFT-e data distributions, and examples of AGC triggering methods were discussed in Sections 6.3.1 and 6.3.2. The AGC is found to improve the Gaussianity of data in most, but not all investigated telescopes and frequency sub-bands. Results from between these sections are not always complimentary. Behaviour of Cambridge's 1350.4 MHz-centred sub-band post-AGC triggering suggests that for sub-bands heavily contaminated with RFI, regular AGC triggering may be necessary in order to keep data in a Gaussian distribution. However, results from Section 6.3.1 indicate that for such environments, AGC triggering may not necessarily redistribute data optimally after one attempt, and so some degree of manual checking may be necessary. This could prove inconvenient for long-duration observations (or commensal observations, e.g., during FRB searches) if some way to automate both the

triggering process *and* verification that optimal data distributions have been achieved is not implemented.

A further inconvenience is presented by the method for manual AGC triggering. A single command-line command is called, which triggers AGC rescaling for all telescopes, sub-bands and polarisations simultaneously. No method for rescaling individual data streams currently exists[20], so inconsistencies between AGC triggers could, if the process is automated, result in a loop whereby some data streams are never optimally distributed. This would impact the overall sensitivities of affected telescopes. It is also noted that in the presence of strong RFI, a fully functional AGC may still adjust the dynamic range of data such that astrophysical signals are confined to a single bin and sensitivity may be lost.

LOFT-e is hardware constrained and cannot receive all sub-bands for all baseline boards at any given time. As data distributions appear to remain consistent for certain sub-bands (e.g., 1414.4 MHz)[21] for a significant amount of time after AGC triggering, one way to circumvent the above issues may be to preferentially avoid e-MERLIN's most RFI-contaminated sub-bands when configuring LOFT-e, and then to manually activate the AGC and obtain decent data distributions. Multiple triggers may be necessary to fully optimize data distributions. A full investigation of AGC application over all permutations of telescope, sub-band and polarisation is advisable if LOFT-e's sensitivity is to be maximised[22].

During these experiments, only one second of data was used to make each histogram provided, and these histograms were offset by time periods of the order minutes to tens of minutes. Depending on RFI fluctuations within a given LOFT-e data file (currently up to 10 minutes in duration) the data may not remain Gaussian throughout this time period. During future commensal LOFT-e observations implementing real-time single-pulse searching, the AGC should not be triggered on a period shorter than the length of time over which data will be single-pulse searched[23].

---

[20]Additionally, this feature is not currently available at the correlator baseline board level.

[21]The 1414.4 MHz sub-band was chosen for these experiments as it is known to have very little RFI contamination. The other sub-bands chosen were chosen here because of their relatively more contaminated RFI environments.

[22]Investigation into baseline-board dependence of AGC behaviour may be necessary.

[23]E.g., the time delay caused by a DM of 3000 pc cm^{-3}.

Chapter 7

Constraining Redshifts of Unlocalised FRBs

This chapter is based on the pre-print entitled "Constraining Redshifts of Unlocalised Fast Radio Bursts" submitted to *Astronomy and Astrophysics* by the author on 04/04/2018 and uploaded to the arXiv on 06/04/2018 (Walker et al., 2018a). Changes reflect referee feedback, and the paper will be resubmitted shortly.

Constraining Redshifts of Unlocalised Fast Radio Bursts

Charles R. H. Walker, Yin-Zhe Ma, and Rene P. Breton

April 6, 2018

7.1   Introduction

As discussed in Section 2.1, as a radio wave travels through an ionised medium it experiences a frequency-dependent index of refraction. For bursts of radio emission from, e.g., pulsars, we measure a time delay on Earth between two observing frequencies given by Equation 2.3 which is proportional to the integrated electron density along the line of sight (DM) (Manchester & Taylor, 1972).

Fast radio bursts (see Section 1.3) have been measured to have DMs in excess of that afforded by the Milky Way alone (see Section 2.1.2). They are therefore assumed to be extragalactic. Extragalactic FRBs from other host galaxies will travel through electrons not only associated with the Milky Way, but also with the IGM, the host

galaxy of the burst, and the local burst environment (Macquart et al., 2015). According to relativistic theories, the observed time delay for these other components will depend on the redshift of the source (Ioka, 2003; McQuinn, 2014; Macquart et al., 2015) and their associated DM factors must be modified.

By studying the relationship between observed DMs and redshifts for many FRBs we may learn about the distributions of electrons which contribute to individual DM components. Statistical studies into potential IGM (e.g., Zhou et al. 2014; McQuinn 2014) and host galaxy (e.g., Xu & Han 2015; Yang et al. 2017) contributions (see Section 2.1.2) to FRB DMs have been undertaken, accounting for, e.g., the structure of the IGM or the locations of FRBs within their hosts. In practice, many FRBs with well constrained redshifts must be accumulated to compare to the results of such studies.

Redshifts may be obtained spectroscopically by association with a host galaxy via interferometry, as has been demonstrated for the Repeater (Tendulkar et al., 2017). However many FRBs have been and will continue to be detected by telescopes with large FoVs, within which their locations may be unconstrained (see Section 3.1.3).

Large FoVs may contain many galaxies. The L-band FoV of a Parkes multibeam receiver beam (HPBW~ 14.4'), for example, may contain $65\,000$ galaxies out to redshift 6 (Scoville et al., 2007)[1]. Molonglo's UTMOST may improve localisation to $15'' \times 15''$ (Caleb et al., 2017), ASKAP's fly's-eye mode has demonstrated localisation of an FRB to $8' \times 8'$ (Bannister et al., 2017) and aims for $< 7''$ localisation when fully operational (Macquart et al., 2010). CHIME will improve localisation for potentially tens of FRBs per day to tens of arcminutes (Ng et al., 2017; The CHIME/FRB Collaboration et al., 2018). The large FoVs and on-sky time available to resolution-limited, single dish and fly's-eye surveys suggest that many FRBs will be discovered to worse than sub-arcsecond accuracy and will never be unambiguously associated with a single host.

Standard calculation of redshifts for unlocalised FRBs often involves simplification of the theoretical relationship between redshift and the IGM component of DM (described in, e.g., Ioka 2003) assuming no host DM component (e.g., Petroff et al. 2016) or a fixed host DM component (e.g., Lorimer et al. 2007; Thornton et al. 2013). The associated uncertainties due to, e.g., position of the source in its host galaxy, or density variations in the IGM (see, e.g., McQuinn 2014) may not be considered.

The purposes of this work are as follows:

---

[1]As expected by COSMOS 2, to a limiting magnitude $I_{AB}(10\sigma) < 27.5$.

1. To provide a framework to perform physically motivated DM-redshift analysis for large numbers of unlocalised FRBs accounting for uncertainties such as (a) potential host contributions to DM (following, e.g., Xu & Han 2015; Yang et al. 2017); (b) the cosmic variance of IGM structure (following, e.g., McQuinn 2014); (c) FRB progenitor evolution across cosmic time (following, e.g., Zhou et al. 2014).

2. To demonstrate this framework via application, using physically motivated models, and to illustrate the effects of different progenitor scenarios on the probability distribution functions of observed dispersion measures and redshifts.

3. To discuss the implications of our models, to provide constrained redshift values for currently unlocalised FRBs and to identify opportunities for DM-redshift analysis during future FRB research.

In Section 7.2 we review the components of extragalactic DM which contribute to the observed time delay between two frequencies observed on Earth for an FRB, and review their respective redshift dependences. We discuss our choice of physically-motivated models for their probability distribution functions (PDFs) when observed in large numbers. In Section 7.3 we present a framework for the combination of these PDFs, in order to retrieve the overall likelihood of an FRB of given redshift having a particular excess DM, or vice versa. We demonstrate this method using our models. Section 7.4 discusses the implications of our findings.

Unless otherwise stated, in this paper we adopt a spatially flat ΛCDM cosmological model with parameters fixed to the best-fit values of Planck Collaboration et al. 2016b: $\Omega_{\rm b} = 0.048$, $\Omega_{\rm m} = 0.309$, $\Omega_{\Lambda} = 0.691$, $n_{\rm s} = 0.9608$, $\sigma_8 = 0.815$, $H_0 = 67.3$.

7.2 Contributions to FRB dispersion measures

From Equation 2.3, the observed time delay $\Delta t_{\rm obs}$ between two frequencies $\nu_{1,\rm obs}$ and $\nu_{2,\rm obs}$ is a function of dispersion measure $\rm DM_{obs}$ which consists of multiple components. Combining Equations 2.4 and 2.5, the complete deconstruction these components is:

$$\rm DM_{obs} = DM_{MW} + DM_{exc} = DM_{MW} + DM_{IGM} + DM_{host} + DM_{local}, \qquad (7.1)$$

where the components represent contributions from ionised electrons along the FRB light path due to the Milky Way (DM_{MW}), IGM (DM_{IGM}), the FRB host galaxy (DM_{host}) and the FRB local environment (DM_{local}) (Macquart et al., 2015). In this work we refer to the DM component in excess of DM_{MW} as DM_{exc}.

7.2.1   The Milky Way and local burst environment

Methods and models used to estimate DM_{MW} are introduced in Section 2.1.2. In order to demonstrate our framework, in the following sections we assume all FRBs simulated for our models have had correct DM_{MW} components discarded. For discussion on DM_{MW} modelling, see Section 7.4).

The DM contribution arising from an FRB's local environment DM_{local} will depend on its progenitor. Potential FRB progenitor scenarios are discussed in Section 1.3.2. Thus far only the Repeater has been localised, allowing more likely candidate progenitors to be identified (see Sections 1.3.3 and 2.1.2). Analysis by Tendulkar et al. (2017) suggests that the Repeater's DM_{local} is dominated by DM_{host}, however this may not be the case for all FRBs. Further examples of localised FRBs are necessary to accurately model this component. For the remainder of this chapter we assume $DM_{local} = 0$.

7.2.2   The intergalactic medium

As discussed in 2.1.2, while $\sim 40\%$ of the Universe's baryonic matter budget is easily observed, the remaining portion is less easily detectable, and only recently are its sources being identified. As FRBs incur their DM_{IGM} component during propagation through the IGM's ionised electrons, they serve as a potential probe of this matter (McQuinn, 2014).

The DM_{IGM} component relates to the observed time delay $\Delta t_{obs,IGM}$ measured between two observed frequencies $\nu_{1,obs}$ and $\nu_{2,obs}$ acquired during propagation through the intergalactic medium, and is therefore dependent on the redshift of the source, z_s, i.e. from Equation 2.3,

$$\Delta t_{obs,IGM} \propto DM_{IGM}(z_s) \left(\frac{1}{\nu_{1,obs}^2} - \frac{1}{\nu_{2,obs}^2} \right). \tag{7.2}$$

Standard methods for analysis of FRB redshifts of DM_{IGM} may discard DM_{host} (see, e.g., Petroff et al. 2016) or subtract a fixed value (e.g., $100 \, \text{pc} \, \text{cm}^{-3}$ Lorimer et al. 2007; Thornton et al. 2013) and then simplify the theoretical relationship (see,

e.g., Ioka 2003) between redshift and the FRB's observed excess dispersion measure to:

$$z_{\rm s} = ({\rm DM_{obs}} - {\rm DM_{MW}})/1200 \, {\rm pc \, cm^{-3}}. \tag{7.3}$$

Ioka (2003) assume homogeneously distributed, 100 % fractionally ionised HI and Helium[2]. Out to $z_{\rm s} \sim 2$, this method recovers the mean value of ${\rm DM_{IGM}}$ for a collection of FRBs originating at a given $z_{\rm s}$ to 2% accuracy (Petroff et al., 2016).

However, variability in the number of collapsed systems, e.g., galactic halos, which lie along the propagation path of any observing sightline will result in a variance around this mean for an observed FRB population (as dicussed by, e.g., Ioka 2003; McQuinn 2014). To obtain more realistic DM-redshift constraints it is necessary to account for this variance.

(McQuinn, 2014) have previously simulated probability distribution functions describing the likelihood of an FRB originating at redshift $z_{\rm s}$ contributing an IGM component ${\rm DM_{IGM}}$ to the total observed ${\rm DM_{obs}}$. To demonstrate our framework, following McQuinn 2014 we use a Gaussian approximation to model this probability distribution function, which we denote $P({\rm DM_{IGM}}|\, z_{\rm s})$:

$$P({\rm DM_{IGM}}|\, z_{\rm s}) = \frac{1}{\sqrt{2\pi}\sigma_{\rm DM}} \exp\left(-\frac{({\rm DM} - {\rm DM_{IGM}}(z_{\rm s}))^2}{2\sigma_{\rm DM}^2}\right), \tag{7.4}$$

with redshift-evolving mean:

$$\mathrm{DM_{IGM}}(z_{\rm s}) = \int_0^{\chi(z_{\rm s})} \mathrm{d}\chi \, \frac{n_{\rm e}(z)}{(1+z)^2}, \tag{7.5}$$

where the $n_{\rm e}(z)$ is the 3-dimensional electron number density at redshift z; and with redshift-evolving variance:

$$\sigma_{\rm DM}^2({\rm DM}, z_{\rm s}) = \int_0^{\chi_{\rm s}} \mathrm{d}\chi (1+z_{\rm s})^2 \bar{n}_{\rm e}^2(0) \int \frac{\mathrm{d}^2 k_\perp}{(2\pi)^2} P_{\rm e}(k_\perp, z_{\rm s}), \tag{7.6}$$

where $\bar{n}_{\rm e}(0)$ is the mean electron density at redshift 0, χ is the comoving distance at redshift z, $\mathrm{d}\chi = c \, \mathrm{d}z/H(z)$ where c is the speed of light and $H(z)$ is the Hubble factor at redshift z, $P_{\rm e}(k,z)$ is the matter power spectrum at redshift z for wavenumber k, and $k_\perp$ is the perpendicular component. For a detailed breakdown of the forms of these

---

[2]The true ratio of ionised to total electrons may be between 0.86-1 (see Appendix A.2). Therefore this assumption should be reasonable after the epoch of reionisation, which is believed to finish $\sim z = 6$ (Fan et al., 2006).

equations we use in our work see Appendix A.3.

We make use of the matter power spectrum from public code CAMB[3], with *Planck* best-fitting cosmological parameters (Planck Collaboration et al., 2016b).

7.2.3   The host galaxy

The host galaxy contribution DM_{host} to a total dispersion measure associated with a source has previously been investigated and found to be dependent on both its host galaxy type and location of the source within the galaxy. Ioka (2003) explore the host galaxy component for GRBs, finding it may contribute up to 10^5 pc cm^{-3} if GRBs originate in star-forming regions. Xu & Han (2015) evaluate FRB DM_{host} for spiral and elliptical galaxies in the rest frame by scaling the NE2001 model and find that they may contribute up to a few thousands or tens of pc cm^{-3} respectively. Yang et al. (2017) estimate the mean host contribution for 21 observed FRBs to be ~ 270 pc cm^{-3}.

As discussed by, e.g., Ioka 2003; Zhou et al. 2014; Macquart et al. 2015, the time delay Δt_{obs} observed on Earth associated with $DM_{host,r}$, the DM contribution to DM_{obs} of an FRB by its host galaxy (at redshift z_s) in its rest frame is dependent on:

1. The relationship between the frequency of the radio wave as emitted in the rest frame of a receding galaxy ν_r and the observed frequency ν_{obs}:

$$\frac{\nu_r}{\nu_{obs}} = (1 + z_s), \tag{7.7}$$

2. The relationship between a time delay measured in the rest frame of the galaxy $\Delta t_{r,host}$ and the observed time delay $\Delta t_{obs,host}$ due to cosmological time dilation:

$$\frac{\Delta t_{obs,host}}{\Delta t_r} = (1 + z_s). \tag{7.8}$$

so that when included in Equation 2.3, the result is an observed time delay $\Delta t_{obs,host}$ between two observed frequencies $\nu_{1,obs}$ and $\nu_{2,obs}$ which scales as $1/(1 + z_s)$ of that it would contribute at $z_s = 0$, i.e.:

$$\Delta t_{obs,host} \propto DM_{host}(z_s) \left(\frac{1}{\nu_{1,obs}^2} - \frac{1}{\nu_{2,obs}^2} \right), \tag{7.9}$$

where

$$\text{DM}_{\text{host}}(z_{\text{s}}) = \frac{\text{DM}_{\text{host,r}}}{1+z_{\text{s}}}, \tag{7.10}$$

represents the time delay due to the rest frame dispersion measure $\text{DM}_{\text{host,r}}$. By assuming that galaxies are static and do not evolve with redshift[4] and by substituting $\text{DM}_{\text{host}} = U(x)$ and $\text{DM}_{\text{host,r}} = x$:

$$P(U)\text{d}U = P(x)\text{d}x, \tag{7.11}$$

it follows that:

$$P(U) = P(x)\frac{\text{d}U}{\text{d}x} = (1+z_{\text{s}})P(x), \tag{7.12}$$

where $U(x) = x/(1+z_{\text{s}})$, $\text{d}U/\text{d}x = 1/(1+z_{\text{s}})$ and:

$$P(\text{DM}_{\text{host}}|z = z_{\text{s}}) = (1+z_{\text{s}})P((1+z_{\text{s}})\text{DM}_{\text{host}}|z=z_0), \tag{7.13}$$

where $z_0 = 0$, and where $\text{P}(\text{DM}_{\text{host}}|z = z_{\text{s}})$ is the probability distribution function (hereafter PDF) describing the likelihood that the host galaxy of an FRB at redshift z_{s} will contribute a host component to DM_{obs} of DM_{host}.

By simulating a host galaxy electron distribution model in the rest frame and populating it with FRBs, PDFs describing the likelihood of a host galaxy at redshift z_{s} contributing a particular DM to the total observed DM may be built for different progenitor models. To demonstrate our framework we simulate simplified electron distribution models for galaxy classes (spiral and elliptical) which we populate with FRBs according to various progenitor scenarios. By integrating electron densities along sightlines we obtain rest-frame DM probability densities which we collect into redshift-scalable PDFs. Our models are described below.

7.2.3.1 Spiral galaxy model

Previous simulations of electron distributions for generic spiral galaxies used in FRB analysis include those by Xu & Han (2015), who scale the Milky Way Cordes & Lazio (2002)'s NE2001 model. However, as simpler models (see Schnitzeler 2012 for a review) have proven sufficient for reproducing the DMs of Milky Way pulsars wih known parallax distances to within a factor 1.5-2, we choose to use a simpler model to

---

[4]Galaxies grow and merge over time, and we do not yet know the progenitors of FRBs. Therefore this is a simplification. Accounting for such evolution could be a source of future work.

Table 7.1: Spiral galaxy electron distribution model parameters (see Gómez et al. (2001); Schnitzeler (2012); Kalberla & Dedes (2008)).

Parameter	Value
$h_{r,1}$	30.4 kpc
$h_{r,2}$	1.5 kpc
$h_{z,1}$	1.07 kpc
$h_{z,2}$	0.050 kpc
n_1	2.03×10^{-2} cm^{-3}
n_2	0.71×10^{-2} cm^{-3}
r_f	17.5 kpc
r_n	3.15 kpc

demonstrate our framework.

We use a model which is the combination of two components:

1. The GBCa model (as denoted by Schnitzeler 2012) which was developed using 109 pulsars with DM-independently determined distances Gómez et al. (2001). This model is a sum of two ellipsoids made from decaying exponentials and describes the inner Milky Way Galactic disk, where the majority of the Milky Way's stellar mass resides. The model is well constrained within 4-12 kpc of the Galactic center. We make the assumption that the model is acceptable within these central 4 kpc for the demonstrative purposes of this work.

2. A model which traces neutral Hydrogen (HI) developed by Kalberla & Dedes 2008. This model is a single decaying exponential and approximates the average mid-plane volume density of the Milky Way disk between 7-35 kpc.

We combine these two components into a smoothly varying function:

$$n_{\text{e,spiral}}(r,h) = \frac{1}{1 + \exp((r - r_f)/r_n)} \sum_{i=1}^{2} n_i \frac{\exp(-r/h_{r,i})}{\exp(-r_\odot/h_{r,i})} \exp\left(\frac{-h}{h_{z,i}}\right), \qquad (7.14)$$

where cylindrical coordinates r and h denote the radial distance from the Galactic center and height above the Galactic plane respectively. All relevant parameters are provided in Table 7.1. To reconcile the two components, the HI model is multiplied by an ionization fraction of $\sim 15\%$.

Our $n_{\text{e,spiral}}(r,h)$ model and its two components are presented in Figure 7.1. At a radius ~ 50 kpc this model reaches an electron density value consistent with the IGM
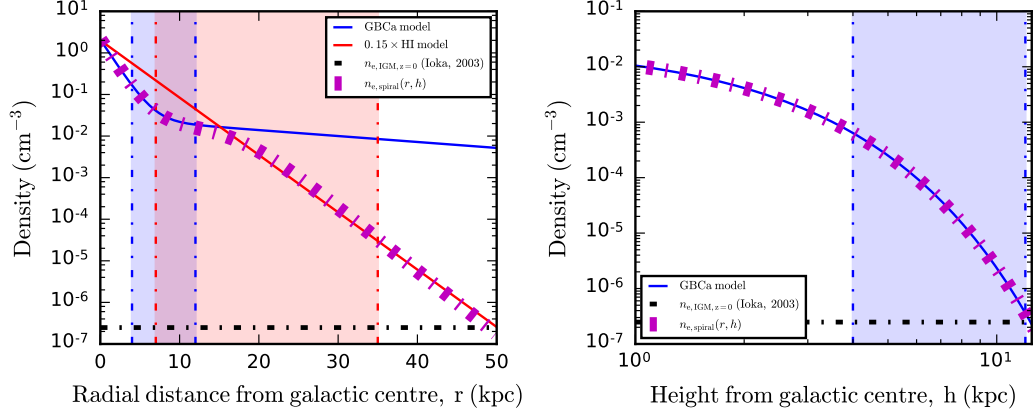
Figure 7.1: Radial and height cross sections of our spiral galaxy electron density model $n_{\text{e,spiral}}(r,h)$ (magenta, dashed). Constituent components of the model tracing the inner MW Galactic disk (Gómez et al., 2001; Schnitzeler, 2012) (blue lines), and an ionised outer HI disk model (Kalberla & Dedes, 2008) are shown (red lines). Shaded (blue/red) regions highlight areas in which the inner/outer profiles are considered well-constrained. The horizontal black dashed line indicates the approximate electron density for the IGM at $z_{\text{s}} = 0$ Ioka (2003) .

at redshift zero, as estimated by Ioka (2003).

7.2.3.2 Elliptical galaxy model

To model an elliptical galaxy electron distribution we follow Mamon & Łokas (2005), using a β model:

$$n_{\text{elliptical}}(r) = \rho_0 \left(1 + \left(\frac{r}{r_{\text{c}}} \right)^2 \right)^{-3\beta_{\text{g}}/2}, \qquad (7.15)$$

(Brown & Bregman, 2001) which traces the density of hot gas within the virial radius of an elliptical galaxy as a function of radius r from its center, central density ρ_0, $\beta_{\text{g}} = 0.5$, and r_{c}:

$$r_{\text{c}} \simeq \frac{R_{\text{e}}}{q}, \qquad (7.16)$$

where $q = 10$ and R_{e} is the effective galaxy radius (i.e. the radius of half-projected light). We choose to model a galaxy of mass $5 \times 10^{10} \, \text{M}_\odot$ with $\rho_0 \simeq 6 \times 10^7 \, \text{M}_\odot \, \text{kpc}^{-3}$ and $R_{\text{e}} = 3.2$ kpc. The relationships between total galaxy mass, central density and effective radius for the β model are provided by Mamon & Łokas (2005). Assuming that this gas may be used as a tracer of electrons, we multiply it by an ionisation fraction (0.04) chosen so that the model approaches the local IGM electron density at
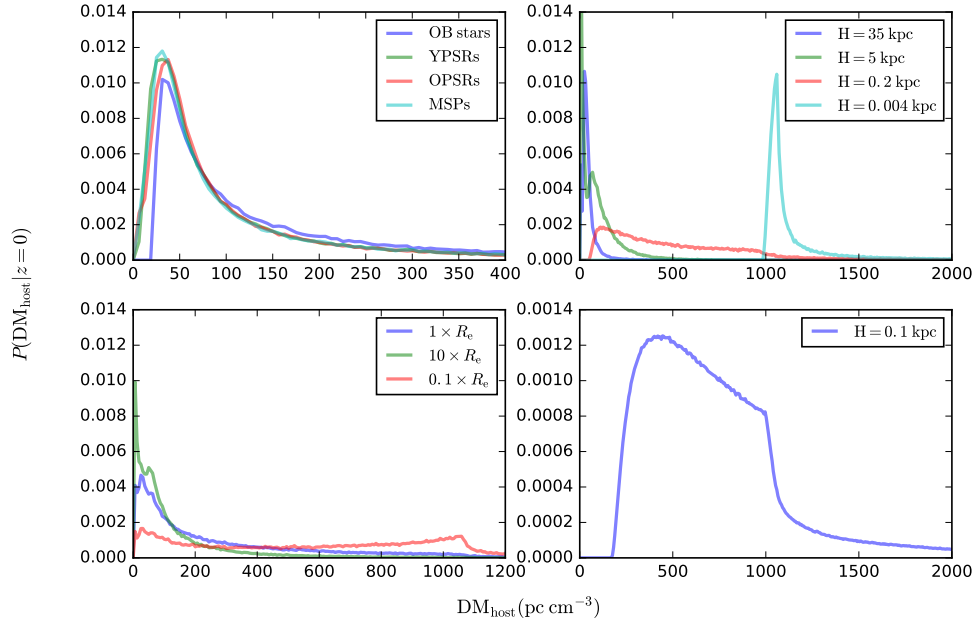
Figure 7.2: Rest-frame PDFs of relative likelihood against DM_{host} for multiple progenitor scenarios according to our models. Sources were generated at different x, y, z coordinates and DMs were calculated for observers on the galaxy boundary at random orientations. From top left to bottom right sources follow: stellar distributions (OB stars, YPSRs, OPSRs, MSPs) in a spiral galaxy; homogeneous distributions (within spheres of radius H kpc) in a spiral galaxy; elliptical distributions (confined to multiples of the galaxy effective radius, $R_e = 3.2$ kpc) in an elliptical galaxy; homogeneous distributions (within spheres of radius H kpc) in an elliptical galaxy.

the galaxy's virial radius (~ 70 kpc).

7.2.3.3   FRB Distributions Within Galaxies

The spacial distributions of FRBs within host galaxies will rely on their progenitors. We populate our host galaxy models with FRBs according to various spacial distributions and calculate DMs for outside observers viewing from random orientations. The resulting $P(DM_{host}|z_s = 0)$ PDFs are shown in Figure 7.2. Our modelling choices are discussed below.

Stellar distributions in spiral galaxies
Stellar populations and phenomena potentially associated with FRBs include massive progenitor stars and superluminous supernovae (Tendulkar et al., 2017), outbursting

Table 7.2: Parameters used in pulsar distribution models (Yusifov & Küçük, 2004; Sun & Han, 2004; Maíz-Apellániz, 2001).

Parameter	Value
σ	7.5 kpc
$r_\odot$	8.5 kpc
r_1	0.55 ± 0.10 kpc
A_{PSR}	1.64 ± 0.11
a	1.64 ± 0.11
b	4.01 ± 0.24
h_{OB}	63 pc
h_0	49×10^{-3} kpc
σ_{PSR}	282 km s^{-1}
h_{OPSR}	0.4 kpc
$t\,(\mathrm{YPSR})$	2 Myr
$t\,(\mathrm{OPSR})$	1 Gyr
h_{MSP}	0.5 kpc

magnetars (Pen & Connor, 2015), supergiant pulses from pulsars (Cordes & Wasserman, 2016), collapsing (Falcke & Rezzolla, 2014) or merging (Totani, 2013) neutron stars, and pulsar companions (Mottez & Zarka, 2014). Such progenitors should follow the spacial distributions of young stellar populations or neutron stars. We choose to model young stellar populations (OB stars), two age populations of pulsars (separated by a characteristic age of 8 Myrs, which we denote YPSRs and OPSRs) and millisecond pulsars (MSPs) separately to account for displacement due to, e.g., supernova kicks.

Height distributions as a function of height above the galactic plane ($H_*(h)$, where * denotes the model type) for OB stars, YPSRs and OPSRs were modelled following Maíz-Apellániz (2001); Sun & Han (2004):

$$H_{\mathrm{OB}}(h) = \frac{1}{\sqrt{2\pi}h_{\mathrm{OB}}} \exp\left(-\frac{h^2}{2h_{\mathrm{OB}}^2}\right), \tag{7.17}$$

$$H_{\mathrm{YPSR}}(h) = \frac{2}{\sqrt{2\pi}h_{\mathrm{PSR}}} \exp\left(-\frac{h^2}{2h_{\mathrm{PSR}}^2}\right), \tag{7.18}$$

$$H_{\mathrm{OPSR}}(h) = \frac{2}{\sqrt{2\pi}h_{\mathrm{PSR}}} \exp\left(-\frac{h^2}{2h_{\mathrm{PSR}}^2}\right) + \frac{1}{h_{\mathrm{OPSR}}} \exp\left(\frac{|h|}{h_{\mathrm{OPSR}}}\right), \tag{7.19}$$

with scale height: $h_{\text{PSR}} = h_0 + \sigma_{\text{PSR}} t$. The height distribution for MSPs was modelled after Lorimer (2013):

$$H_{\text{MSP}}(h) = \frac{1}{h_{\text{MSP}}} \exp\left(-\frac{h}{h_{\text{MSP}}}\right). \tag{7.20}$$

Parameters are provided in Table 7.2.

Radial distributions as a function of radius from the galactic center ($R_*(r)$, where * denotes the model type) for YPSRs and OPSRs were modelled with a gamma function following Yusifov & Küçük (2004) and with a Gaussian distribution for MSPs and OB stars following Lorimer (2013):

$$R_{\text{YPSR}}(r) = R_{\text{OPSR}}(r) = A_{\text{PSR}} \left(\frac{X}{X_\odot}\right)^a \exp\left(-b\left(\frac{X - X_\odot}{X_\odot}\right)\right), \tag{7.21}$$

where $X = r + r_1$ and $X_\odot = r_\odot + r_1$.

$$R_{\text{OB}}(r) = R_{\text{MSP}}(r) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{r^2}{2\sigma^2}\right). \tag{7.22}$$

Parameters are provided in Table 7.2.

Other distributions

To populate elliptical galaxies we use the stellar mass distribution (and thus the ionised electron distribution) of the galaxy as a tracer for FRB distributions by populating the host with sources distributed following $n_{\text{elliptical}}(r)$, normalised to the number of sources simulated. We also generate homogeneously distributed sources within spheres of variable radius for both spiral and elliptical galaxies. The associated distributions attempt to simulate progenitors in central galactic halos. For smaller limiting radii such distributions could represent FRB mechanisms attributed to AGN, e.g., interactions between relativistic jets and plasma (Romero et al., 2016; Vieyro et al., 2017). Larger limiting radii could be consistent with progenitors located in more extended galactic halos. We note that the most accurate resulting DM distributions for the smallest limiting radii require electron distribution models with well-defined central profiles.

7.3   The Excess Electron Model

Here we present the framework for combining the PDFs of each contributing factor to DM_{exc} to retrieve a PDF which describes the likelihood that an FRB at source redshift
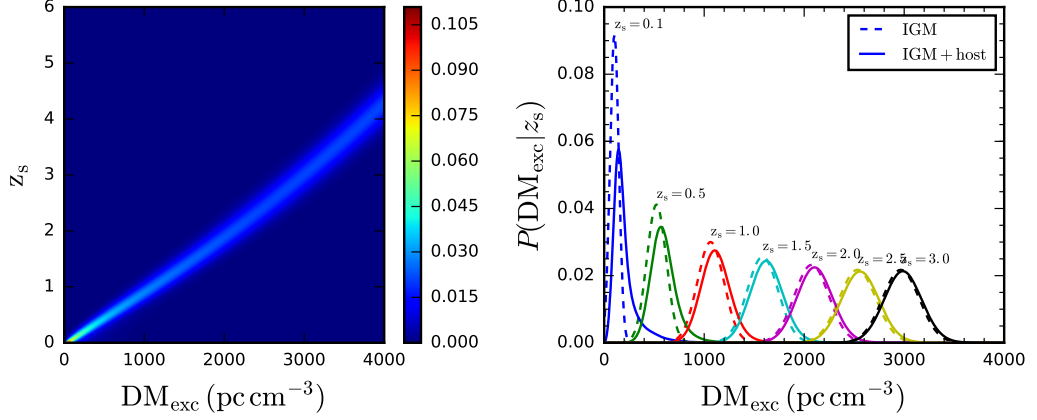
Figure 7.3: Excess electron model for a YPSR-like FRB population in spiral galaxies. Left: Two-dimensional PDF of source redshift z_s against observed excess dispersion measure DM_{exc}. Colour indicates the relative likelihood that a source at z_s will have an observed DM_{exc}. The model is normalised to ensure $\int P(DM_{exc}|z_s)\,dDM = 1$. Right: Projections of the model for different source redshifts (solid lines) compared with our chosen IGM probability distribution $P(DM_{IGM}|z_s)$ (dashed lines). Curves are normalised to unity.

z_s will have an observed extragalactic DM component DM_{exc}, which will result in a measured time delay $\Delta t_{obs,exc}$ between two observing frequencies $\nu_{1,obs}$ and $\nu_{2,obs}$ on Earth. We demonstrate the method by combining our simulated PDFs. The method is applicable to all IGM and host models, and so in the future may be used with more accurate models derived from FRB discoveries. We then apply Bayes' theorem to the model and discuss advantages of doing so.

7.3.1 P(DM|z) method and results

An FRB emitted from a source at redshift z_s will have a DM_{exc} which is the sum of its DM_{IGM} and DM_{host} components. Assume these components to be random independent variables drawn from their respective underlying PDFs, $P(DM_{IGM}|z_s)$ and $P(DM_{host}|z_s)$. The PDF for summed independent random variables is the convolution of the PDFs of their components, therefore:

$$P(DM_{exc}|\,z_s) = P(DM_{IGM}|\,z_s) * P(DM_{host}|\,z_s) \tag{7.23}$$

where $P(DM_{exc}|\,z_s)$ describes the likelihood that an FRB of given source redshift z_s will have an observed extragalactic DM component DM_{exc}. This DM_{exc} will result in a measured time delay $\Delta t_{obs,exc}$ between two observing frequencies ($\nu_{1,obs}$ and $\nu_{2,obs}$)

on Earth. In this work we refer to $P(\mathrm{DM_{exc}}|\,z_s)$ as the excess electron model.

An example excess electron model generated using our IGM and YPSR PDFs is shown in Fig. 7.3. By comparing projections of the model generated for given z_s with $P(\mathrm{DM_{IGM}}|z_s)$, the figure illustrates the influence of the host galaxy on $P(\mathrm{DM_{exc}}|z_s)$ with increasing redshift. At low redshifts the host increases the likelihood that an FRB will have a $\mathrm{DM_{exc}}$ significantly in excess of that predicted by standard analysis which considers only electrons in the IGM. This influence diminishes with increasing redshift, signified by the curves almost converging by $z_s = 3$.

Figure 7.4 illustrates this effect for combinations of our host models. Compared models follow spacial distributions of YPSRs in spiral galaxies, homogeneous distributions close to the galactic centers of spiral galaxies, elliptical distributions confined to within $0.1 \times R_e$ and $1 \times R_e$ where $R_e = 3.2$ kpc is the elliptical galaxy's effective radius, and homogeneous distributions close to the galactic centers of elliptical galaxies. The effect of the host is significant at low z_s and for some distributions (e.g. those where FRBs are confined to areas of large central density) out to $z_s = 3$. For large FRB populations with accurately known redshifts this could be a potentially useful tool for identifying host/progenitor types (see Sect. 7.4). We note, however, that these curves are most accurate for FRB populations consisting of a single host/progenitor and some combination of these models may be more accurate in practice.

7.3.2   P(z|DM) method and results

According to Bayes' theorem (Bayes & Price, 1763; Laplace, 1812):

$$P(z_s|\,\mathrm{DM_{exc}}) = \frac{P(\mathrm{DM_{exc}}|\,z_s)P(z_s)}{P(\mathrm{DM_{exc}})}, \tag{7.24}$$

where the $P(z_s)$ is the assumed prior redshift distribution of FRBs, and $P(\mathrm{DM_{exc}})$ is the probability of observing a DM excess. The resulting PDF $P(z_s|\,\mathrm{DM_{exc}})$ describes the likelihood that an observed FRB with a measured time delay $\Delta t_{\mathrm{obs,exc}}$ between two observing frequencies ($\nu_{1,\mathrm{obs}}$ and $\nu_{2,\mathrm{obs}}$) on Earth due to a DM excess $\mathrm{DM_{exc}}$ will have a source redshift z_s according to the IGM and host models used in $P(\mathrm{DM_{exc}}|\,z_s)$. For practical application this model should be normalised so that $\int P(z_s|\,\mathrm{DM_{host}})\,\mathrm{d}z = 1$.

Projections of $P(z_s|\,\mathrm{DM_{exc}})$ generated using a single host and IGM model for an FRB with given $\mathrm{DM_{exc}}$ values compared against the model generated using the IGM component alone illustrate the influence of the host galaxy on redshift probability. Figure 7.5 demonstrates this effect for our YPSR model, and illustrates the diminishing
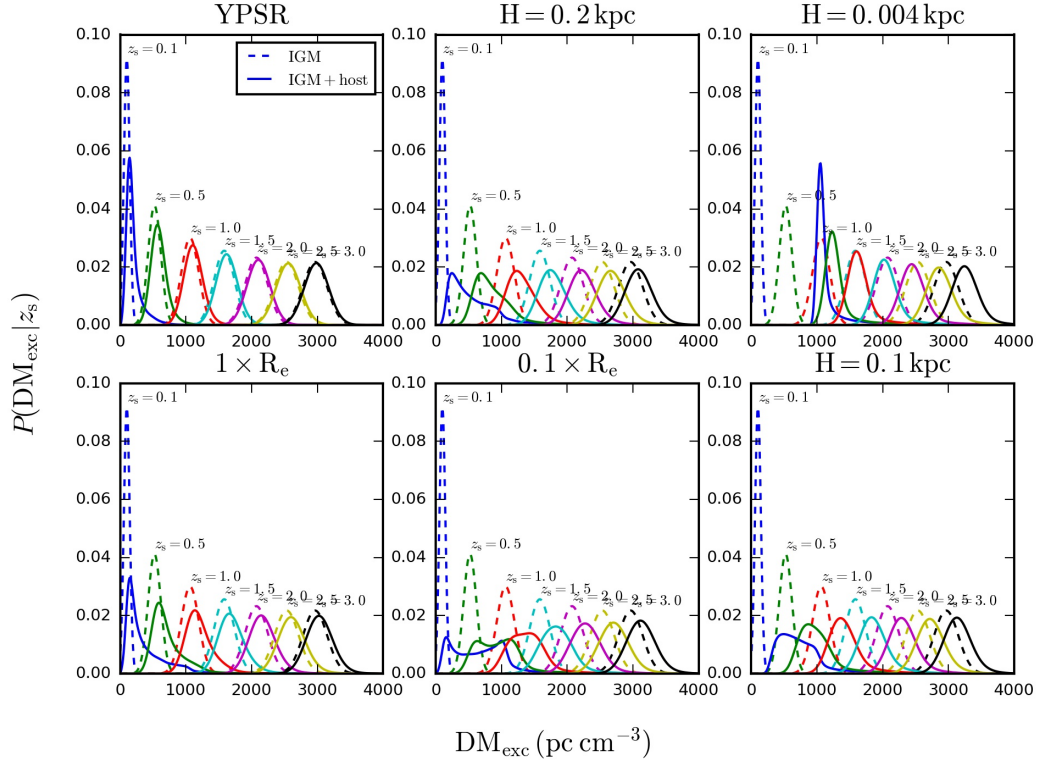
Figure 7.4: Projections of excess electron models generated using several of our $P(\mathrm{DM}_{\mathrm{host}}|z_{\mathrm{s}})$ distributions following the style of Figure 7.3 (right). Top panel (left to right): sources generated from young pulsars in spiral galaxies, and from homogeneous distributions within spheres of radius $H = 0.2$ kpc, $H = 0.004$ kpc from the galactic center. Bottom panel (left to right): sources generated from elliptically distributed FRBs confined to multiples of 1, 0.1 times the galaxy half light radius ($\mathrm{R}_{\mathrm{e}} = 3.2$ kpc), and from a homogeneous distributions within a sphere of radius $H = 0.2$ kpc from the galactic center.
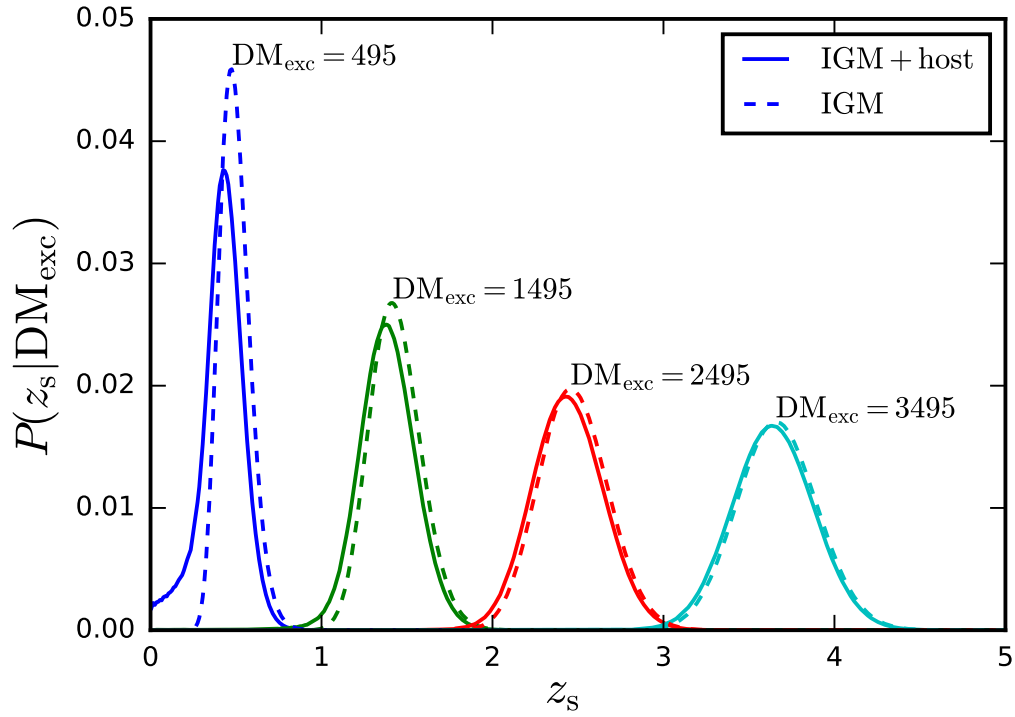
Figure 7.5: Projections of the excess electron model after application of Bayes' theorem assuming a prior redshift distribution $P(z_s) = 1$, and normalisation to $\int P(z_s|DM)\,dz = 1$. Curves indicate the relative likelihood that an FRB with given DM_{exc} originates from a source redshift z_s. Compared curves are generated from a YPSR-like FRB population occupying spiral galaxies, and a population where the host galaxy component is neglected.

influence of the host galaxy on redshift probability with increasing DM_{exc}. Excluding FRBs localised to host galaxies with spectroscopic redshifts measurements, FRB redshifts should therefore be more constrained for those with larger DM_{exc}. We note that best constraints result from accurately defined host progenitors and realistic IGM models.

The factor $P(z_s)$ which arises after application of Bayes' theorem is the assumed prior redshift distribution of FRBs across cosmic time. Setting $P(z_s) = 1$ implicitly assumes that the likelihood of an FRB occurrence at any given redshift is equally probable. As redshift distributions will rely on the phenomenology of their progenitors this is unlikely to be the case. Zhou et al. (2014) have previously considered the redshift distribution of FRBs to follow an Erlang distribution fit:

$$P(z_s; k, \gamma) = \gamma^k z_s^{k-1} e^{-\gamma z_s} / \Gamma(k), \qquad (7.25)$$

with $k = 2$, $\gamma = 1$ which has been associated with GRBs by (Shao et al., 2011). Figure 7.6 illustrates the effect of applying an Erlang distribution prior to the excess electron model generated for an FRB with DM_{exc} equal to that of the Repeater. The prior affects the FRB's potential redshift range, resulting in a decreased likelihood of the FRB originating at a lower redshift.

7.4 Discussion and Conclusion

Our previous sections reviewed the components which contribute to observed FRB dispersion measures and their redshift dependence, and combined physically motivated PDFs for those components. The resulting excess electron model allows more accurate constraints to be placed on unlocalised FRB redshifts than arise from standard practice. This is beneficial as an FRB's observed DM will always be well-defined regardless of whether or not a spectroscopic redshift may be obtained. The aims of our model are threefold:

1. To improve FRB DM-redshift analysis out to $z > 2$ by considering:
   - The redshift-evolving mean in DM_{IGM} along with the variance acquired due to galactic halos.
   - The contribution of the host galaxy to DM_{exc} under different progenitor assumptions.
   - The effects of underlying FRB redshift distributions on their potential DM_{exc} values.
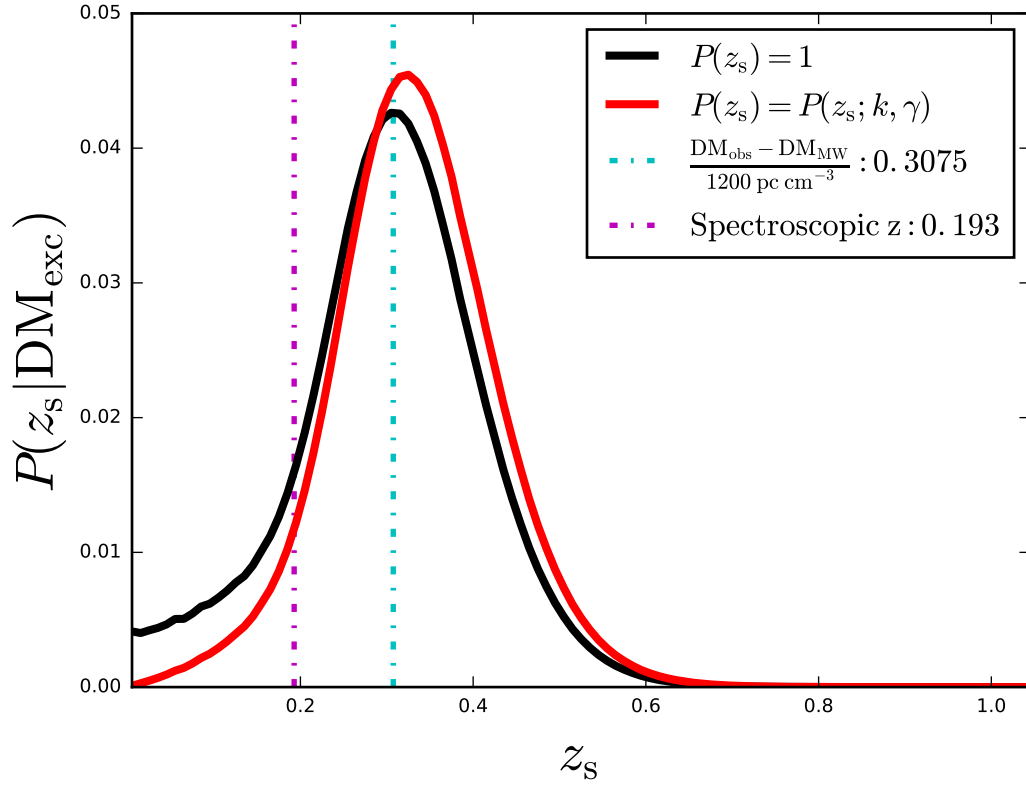
Figure 7.6: Projections of the excess electron model as described for Fig. 7.5 for an FRB with DM equal to that of the repeating FRB 121102. A model generated assuming a homogeneous FRB redshift prior (black line) and a model generated assuming an Erlang redshift prior with $k = 2$, $\gamma = 1$ (red line) are compared. The cyan dash-dotted line indicates the estimated redshift of the FRB as calculated via standard methods. The magenta dashed line indicates the spectroscopic redshift acquired for the host galaxy of FRB 121102.
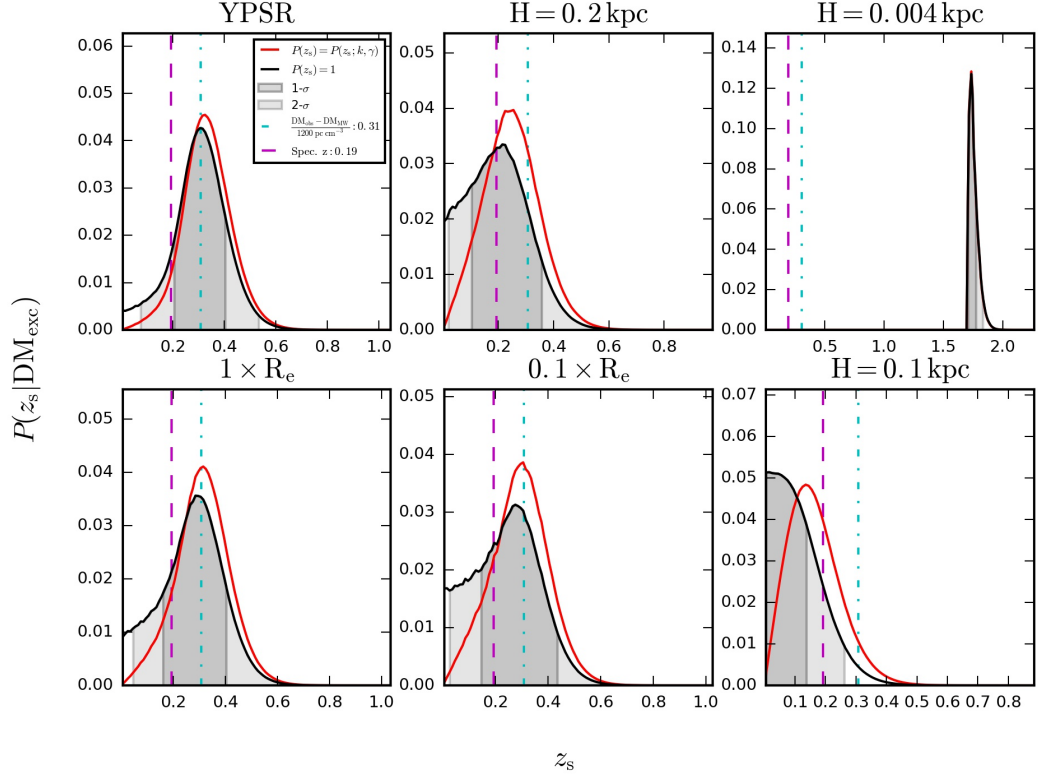
Figure 7.7: Projections of the excess electron model as described for Fig. 7.6 for an FRB with a DM equal to that of repeating FRB 121102. Curves are generated according to progenitors in spiral galaxies (top panel) and elliptical galaxies (bottom panel). Progenitors follow distributions of: top left to bottom right: YPSRs, randomly distributed sources confined to spheres of radius $H = 0.2\,\mathrm{kpc}, 0.004\,\mathrm{kpc}$ around the galactic centers of spiral galaxies, elliptically distributed sources confined to multiples of an elliptical galaxy's half light radius ($R_\mathrm{e} = 3.2$ kpc), randomly distributed sources confined to spheres of radius $H = 0.1$ kpc around the galactic centers of elliptical galaxies. Shaded regions mark bounds drawn from 68% and 99% confidence intervals for the curve generated with a homogeneous redshift prior.

and to illustrate the differences that these physically-motivated simulations have on FRB DM-redshift probability distributions.

2. To place constraints on the redshifts of unlocalised FRBs using their dispersion measures, given the above considerations.

3. To identify potential uses for DM-redshift analysis in future FRB research.

7.4.1   DM component analysis

For all progenitor scenarios we consider, our resulting $P(\mathrm{DM_{host}}|z_\mathrm{s} = 0)$ curves show that FRBs may be dispersed significantly (> 400 pc cm^{-3}, and in some cases in excess of 700 pc cm^{-3}) by the host galaxy in the rest frame (see Fig. 7.2). This result is consistent with previous studies of spiral galaxies (Xu & Han, 2015; Zhou et al., 2014). However our simulations also predict this behaviour for elliptical galaxies, in disagreement with past literature (Xu & Han, 2015). This discrepancy most likely arises due to different modelling methods. Xu & Han (2015) scale a Milky Way-sized elliptical galaxy (constructed by neglecting the thin disk and spiral arm components of the NE2001 model) using H$_\alpha$ as a tracer of ionised gas, whereas we use a β model and confine sources to areas within the galaxy's effective radius.

As demonstrated in Figure 7.4, the relative contribution of the host galaxy to $\mathrm{DM_{exc}}$ is most influential at low z_s and diminishes as z_s increases. This effect, due to cosmological time dilation and recession of the host (Ioka, 2003) takes place regardless of host model used, and has previously been noted to inadequately suppress host galaxy contributions larger than 700 pc cm^{-3} for FRBs with $z_\mathrm{s} < 2$ (Zhou et al., 2014; Macquart et al., 2015). According to our models, in some cases the host may noticeably influence an FRB's potential $\mathrm{DM_{exc}}$ range out to $z_\mathrm{s} \sim 3$ (see, e.g., homogeneously distributed FRBs originating close to galactic centers).

Our chosen models may not accurately represent the origins of FRBs on an individual basis. For example, the Repeater, the only localised source of FRBs to date, originates in a dwarf galaxy (Tendulkar et al., 2017): a scenario which we have not considered. We also note that we hold fixed our spiral and elliptical galaxy masses. More realistic host galaxy mass distributions could be used in the future to improve analysis. As our assumed masses lie closer to the top of the mass range, this may further supress the overall host contribution to $P(\mathrm{DM_{exc}}|z_\mathrm{s})$ relative to that of the IGM. Future models would ideally also incorporate a $\mathrm{DM_{local}}$ component.

7.4.2 Redshift constraints for unlocalised FRBs

Table A.1 presents constrained source redshifts for every catalogued FRB to date, with bounds drawn from their 99% confidence intervals. The table focuses on scenarios which highlight the difference that hosts can make to an FRB's likely redshift range. In all cases a prior redshift distribution of $P(z) = 1$ has been assumed, however the true cosmological FRB redshift distribution will depend on FRB phenomenology. Scenarios chosen are:

1. Distributions neglecting a DM_{host} component. At redshifts $z_s < 2$, these essentially recover values obtained using standard practices.
2. Sources tracing young pulsar populations in Milky Way-like spiral galaxies.
3. Sources confined to galactrocentric spheres of radius $0.2\,kpc$ in Milky Way-like spiral galaxies.
4. Sources confined to galactrocentric spheres of radius $0.1\,kpc$ in $5 \times 10^{10}\,M_\odot$ elliptical galaxies.
5. Sources tracing the gas profile of a $5 \times 10^{10}\,M_\odot$ elliptical galaxy, confined to an area within its $3.2\,kpc$ effective radius.

We provide our models and a Jupyter notebook with this work so that our results may be reproduced, and so that others may obtain redshifts for FRBs with our models or substitute their own $P(DM_{IGM}|z_s)$, $P(DM_{host}|z = z_0)$ and $P(z)$ distributions. We propose that such techniques be considered for general use during future FRB analysis in order to better constrain the redshifts of unlocalised FRBs.

7.4.3 Potential uses for DM-redshift analysis

Here we qualitatively identify some uses for the DM-redshift analysis framework we have presented in this work.

1. **Progenitor distinction**: As shown by Figure 7.4, the choice of host galaxy and progenitor distribution may noticeably affect the likelihood of an observed FRB at given redshift z_s having a particular DM_{exc}. When many FRBs of a given redshift are obtained, this may allow the observed excess DM distribution to be compared to simulated $P(DM_{exc}|z_s)$ curves in order to rule out different progenitor scenarios. We note that this technique will require accurate simulated models of both potential host galaxy electron distributions and spacial distributions of potential progenitor sources within them.
2. **Host candidate rejection**: An observed FRB will always have a well defined

DM_{obs}. By subtracting the DM_{MW} component and selecting a progenitor model and prior redshift distribution, a $P(DM_{exc}|z_s)$ distribution may be generated and the likely redshift range of the FRB may be constrained to a given statistical significance. Host galaxies within the localisation radius of the FRB which lie outside of this redshift range may then be rejected. This requires accurate estimation of the DM_{MW} component. Potential errors from current Milky Way electron density models are discussed in Section 2.1.2. If the prior redshift distribution of FRBs $P(z_s)$ suppresses differences between progenitor models (see, e.g., Figure 7.7) then this technique may be particularly effective.

3. **Milky Way uncertainty mitigation**: By observing a small area of sky, uncertainties due to sightline variations in DM_{MW} would be mitigated. Over a small enough area, the Milky Way contribution to DM_{obs} could be considered uniform, and would result in a constant systematic offset in the observed FRB DM_{exc} distribution in a particular direction for all observed FRBs. Current and future FRB searches with the ability to piggyback deep extragalactic surveys (e.g., V-FASTR, MEERTRAP, LOFT-e Wayth et al. 2011; Burke-Spolaor et al. 2016; Stappers 2016; Walker et al. 2018b) could potentially provide such sets of FRBs. Assuming FRBs have no preferential latitude dependence (Bhandari et al., 2018) this approach would yield similar numbers of events in any pointing direction.

4. **A new Milky Way DM model**: By considering the above approach, sufficient numbers of pointings by deep extragalactic FRB searches over large sky areas, collecting sufficiently large number of FRBs would allow for disentanglement of DM_{exc} and DM_{MW}. Assuming that the true underlying $P(DM_{exc}|z_s)$ is direction-independent, the observed distribution of DMs for FRBs of given redshifts in different directions could be compared and searched for systematic offsets. Flux-limited surveys probing a small redshift range (e.g., ASKAP, CHIME Macquart et al. 2010; The CHIME/FRB Collaboration et al. 2018) could potentially provide enough FRBs to construct a sky map of DM_{MW}, or to implement corrections to existing models, even with limited localisation. As DM_{MW} must be accurately removed to enable cosmological analysis of future populations of FRBs with sperctroscopically-determined redshifts, this is an area where detecting large numbers of FRBs with limited localisation may be relevant.

5. **IGM-host DM component disentanglement**: $DM_{IGM}(z_s)$ is a function integrated over redshift, while $DM_{host}(z_s)$ is evaluated at a specific source redshift. As such, the two components may theoretically be disentangled. By drawing

an analogy between DM component redshift dependence and the frequency-dependence of CMB/Galactic foreground emission, component separation techniques employed by, e.g., Planck Collaboration et al. 2016a may be of use. The main difference, however, is that even though the redshift dependence is known – which in the CMB analogy would be the observed frequency – the shape of the actual underlying PDFs will not be readily extractable. Component separation will likely require forward modelling using template IGM and host PDFs inspired from realistic models as presented in this paper, or a decomposition into basis functions such as a Gaussian mixture. Accomplishing this will necessitate redshifts to be available.

7.4.4 Concluding remarks

We present a framework for exploration of the statistical relationship between FRB redshifts and dispersion measures which provides the basis for:

1. Assessment of host galaxy contributions to FRB dispersion measures using physically-motivated models. We demonstrate this, finding that all of our host models may contribute large DMs ($> 400 \, \mathrm{pc \, cm^{-3}}$) in the rest frame. Through application of relativistic effects in an expanding Universe, we demonstrate that the influence of $\mathrm{DM_{host}}$ diminishes with increasing redshift, but that for extreme progenitor scenarios where FRBs originate close to galactic centers, the $\mathrm{DM_{host}}$ component may still contribute significantly to overall $P(\mathrm{DM_{exc}}|z_\mathrm{s})$ profiles out to $z_\mathrm{s} = 3$.

2. Constraining the redshifts of unlocalised FRBs more rigorously than is current standard practice. By consulting $P(z_\mathrm{s}|\mathrm{DM})$ probability distributions generated using physically-motivated progenitor models, this technique may allow host candidate rejection for unlocalised FRBs, and may allow insight into FRB progenitors to be drawn from large observed populations. A repository containing Python code and examples of the DM-redshift analysis demonstrated in this work is provided at: `https://doi.org/10.5281/zenodo.1209920`.

3. Progenitor distinction, host galaxy rejection, Milky Way n_e uncertainty mitigation and the disentanglement of $\mathrm{DM_{obs}}$ components. Assuming a direction-independent $P(\mathrm{DM_{exc}}|z_\mathrm{s})$, the $\mathrm{DM_{MW}}$ contribution to $\mathrm{DM_{obs}}$ could be extracted by comparing histogrammed DM distributions from flux-limited surveys (e.g. CHIME) on different sightlines and looking for systematic offsets in their profiles. For surveys providing large numbers of FRBs with a small enough redshift range, this technique may not require redshift measurements, thus increasing the

usefulness of unlocalised FRBs. It may also be possible to separate DM_{IGM} and DM_{host} using their respective redshift dependences.

Chapter 8

Concluding Remarks

Chapter 1 of this thesis introduce the phenomena of pulsars and fast radio transients which are detected as bursts of radio emission by telescopes on Earth. Chapters 2 and 3 review significant physical effects undergone by these signals during their propagation through the ionised matter which is prevalent in space, and then summarise techniques of relevance which are used to detect them. The major accomplishments in this thesis are presented in Chapters 4 through 7. These may be summarised as the development of: a new observing mode for the UK's e-MERLIN interferometer, enabling the study of pulsars and fast radio transient phenomena; and a probabilistic framework enabling study of the relationship between a cosmological fast radio transient source's redshift and the measure of dispersion its signal undergoes during propagation through the Universe. These accomplishments and their implications for future work are briefly reviewed below.

8.1   On the e-MERLIN observing backend, LOFT-e

Chapter 4 introduces the necessary data, telescope metadata, VDIF export metadata and observing schedule metadata which are required to create a functioning high time resolution observing backend for the the e-MERLIN interferometer. The author has implemented a backend which streams this data and metadata to three computing nodes via signal paths shown in Figure 4.4. The resulting system may capture up to 12×64 MHz bandwidth, dual polarisation, 2-bit voltage data streams from the e-MERLIN correlator and its associated telescopes. The system enables:

1. **Multi-band data capture**: Data from e-MERLIN's multiple observing bands

may be captured. These bands are: L-band (1.3GHz - 1.8 GHz), C-band (4 GHz - 8 GHz), and K-band (22 GHz - 24 GHz).

2. **Flexible configuration**: An observer may select which data are to be captured from all possible permutations of:

   - **Telescope**: All of e-MERLIN's seven telescopes are compatable with the developed system. These telescopes are the Cambridge, Darnhall, Defford, Knockin, Pickmere, Mark II and Lovell telescopes.
   - **Frequency sub-band**: During a given observation, e-MERLIN's 512 MHz observing bandwidth is split into eight sub-bands. Ideally, any combination of sub-bands may be selected, however this depends on which correlator baseline boards are functional at the time of observation.

3. **Flexible data distribution**: Data may be distributed variably across the system's three computing nodes, `transient`, `compute-0-0`, and `compute-0-1`. This enables evenly distributed data rates across nodes, or flexibility in the event of hardware failure.

The system is named **LOFT-e** (Localisation of Fast Transients with e-MERLIN).

Successful data capture by LOFT-e has been demonstrated, and periodic RFI which occurs to some capacity in data from all telescopes has been identified (e.g., Figure 4.7). Early investigations indicate that this RFI is consistent across sub-bands for a given telescope, indicating a pre-correlator origin. Areas of future improvement to the LOFT-e signal chain have been identified. Such improvement includes removal of unnecessary points of failure in the telescope metadata signal path (see Section 4.6).

8.2   A LOFT-e data processing review

Chapter 5 introduces the LOFT-e digital backends developed by the author. These include:

1. A **high time resolution data capture and observing mode** for e-MERLIN, co-developed with fellow PhD candidate T. W. Scragg.
2. A functional **data processing pipeline** for high time resolution e-MERLIN data.

LOFT-e's observing mode is fully operational. Its development has allowed the author and others to successfully secure observing time for high time resolution radio astronomy during multiple e-MERLIN proposal cycle calls[1]. During development, issues with data capture (e.g., data writing bottlenecks and packet loss, see Section 5.3) have been identified and remedied. A user guide for data aquisition using this observing mode has been provided (see Appendix A.1).

The LOFT-e data processing pipeline is shown in Figure 5.5. The pipeline has been developed to allow:

1. **Filterbanking**: Captured data are filterbanked offline on a data stream by data stream basis. The current standard LOFT-e output data product is a Stokes I sigproc filterbank file of duration ten minutes, with 64 MHz bandwidth, 0.25 MHz spectral resolution and 64 μs temporal resolution.

2. **RFI mitigation**: Captured data may optionally undergo multiple stages of RFI mitigation. These stages include:
    - **Masking**: Files may be masked in the frequency domain on a telescope by telescope basis. A systematic investigation is required to identify persistent narrow-band RFI and develop ideal masks for all permutations of e-MERLIN's telescopes and frequency sub-bands.
    - **Time domain clipping**: An algorithm to supress the periodic RFI described in Chapter 4 has been developed.
    - **Frequency domain MAD flagging**: GPU algorithms enabling frequency domain MAD flagging have been incorporated into the pipeline.

    Tools for implementation of these stages have been developed, and are described in Section 5.2. Improvements to the Gaussianity of data post-RFI mitigation have been demonstrated. These techniques display variable degrees of effectiveness on a telescope and frequency sub-band basis (e.g., figures 5.2.2 and 5.12).

3. **Incoherent beamforming**: Captured data for a given observation may be incoherently beamformed to improve e-MERLIN's sensitivity. Three incoherent beamforming modes have been developed to attempt to mitigate the effects of RFI on data during this process. The modes are:
    - **Mean combination**: Whereby all recorded telescopes are combined, potentially affording up to a $\sqrt{\text{N}}$ sensitivity increase on an individual dish,

---

[1]The e-MERLIN proposal codes utilising LOFT-e which have been awarded observing time as of 22/09/2018 are: CY3219, CY5001, CY5227, CY5228, CY6003, CY6222, CY6226, CY7001, CY7218, CY7219.

where N is the number of telescopes used in the observation.

- **Median outlier rejection**: Whereby the telescope sample furthest from the median is discarded on a sample-by-sample basis, potentially discarding non-Gaussian fluctiatuons in the data and achieving a sensitivity increase closer to that theoretically expected for incoherent beamforming.
- **Median combination**: Whereby the median value of telescope samples is selected on a sample-by-sample basis.

4. **Frequency sub-band combination**: Where multiple frequency sub-bands have been captured for a given telescope and a given observation, they may be combined to potentially afford a $\sqrt{n}$ sensitivity increase, where n is the number of sub-bands combined.

5. **Offline pulsar analysis**: Captured data may be processed using standard pulsar analysis tool packages. The LOFT-e pipeline implements DSPSR as standard.

6. **Offline fast transient analysis**: Captured data may be searched for single pulses using standard tool packages. The LOFT-e pipeline implements AstroAccelerate as standard.

Avenues for future improvement to LOFT-e's data processing pipelines have been identified in Section 5.4. Such improvements include: a fully commensal, real-time pipeline, which requires development of efficient RFI clipping and beamforming software; polarisation analysis, which requires updated filterbanking software which outputs all Stokes data products; coherent beamforming capability to enable increased sensitivity of the instrument; and coherent dedispersion capability to enable the study of the fine structural behaviour of any events observed by e-MERLIN.

8.3   On the study of radio emission with LOFT-e

Chapter 6 reviews early attempts to observe pulsars with e-MERLIN, during which sources of non-Gaussian behaviour in data which hindered pulsar detection (e.g., airport radar, e-MERLIN's AGC) were identified. After implementation of a new AGC control method, pulsars were successfully observed. Sections 6.1.2 and 6.1.3 present, for the first time, LOFT-e observations of multiple pulsars (PSR B0329+54, PSR J1916+1616) at multiple observing frequencies[2], using combinations of all six of e-MERLIN's standard telescopes. Effects consistent with scintillation (see Figure 6.10)

---

[2]Centered on 1350.4 MHz, 1414.4 MHz, 1542.4 MHz and 1670.4 MHz.

and mode switching (see the difference between pulse profiles in Figures 6.11 and 6.12) have been observed for PSR B0329+54. These results demonstrate the system's capability not only to observe pulsars, but also to identify changes in their emission. Potential for future work includes investigating the limits to which LOFT-e may do this for fainter sources, given e-MERLIN's current sensitivity.

The effects of RFI on the sensitivity of different e-MERLIN frequency sub-bands (e.g., Figure 6.5) and different telescopes (e.g., Figure 6.7) to pulsars have been presented. The sub-band centred on 1414.4 MHz is consistently the least contaminated band, demonstrating the highest S/Ns for integrated pulse profiles. Sub-bands centred on 1670.4 MHz and 1542.4 MHz have been shown to suffer from mitigatable RFI. The sub-band centred on 1350.4 MHz has been shown to suffer from severe RFI. These results are not a comprehensive analysis of the sensitivity of e-MERLIN's available telescopes, observing frequencies and sub-bands, and a systematic review will be necessary to identify the bands which will be most effective for pulsar observations or fast radio transient searches.

The RFI mitigation techniques introduced in Chapter 5 have been demonstrated to improve the S/Ns of the integrated pulse profiles of pulsars recovered during LOFT-e observations. Their effectiveness has been shown to vary over telescopes and frequency sub-bands. In some cases, applied techniques supress noise levels enough for pulsars to be identified where they were not visible in raw data (e.g., Cambridge observations at 1542.4 MHz in Figure 6.5). In some cases, RFI mitigation is ineffective and pulse profiles cannot be recovered (e.g., Cambridge telescope observations at 1350.4 MHz in the same figure). The effectiveness of each RFI mitigation stage has been tested by via study of recovered pulse profile S/Ns after each successive mitigation stage. Masking with crudely developed frequency-domain masks is not necessarily an effective way of increasing the S/N of an observation, so further research into persistent narrow-band RFI at the telescope, frequency sub-band and observing band level may be necessary for optimum application of this technique. Time domain RFI clipping proves consistently effective in increasing S/N, and may contribute up to a factor of 2 increase in sensitivity (e.g., Figures 6.6 and 6.8). Frequency domain MAD flagging is consistently ineffective for RFI reduction, implying a less severe threshold (i.e. a larger clipping value) should be used during its implementation.

Frequency sub-band combination has been demonstrated, and shown to afford less than the factor $\sqrt{n}$ sensitivity increase (where n is the number of combined sub-bands) implied by the radiometer equation (e.g., Figure 6.9). This is unsurprising due to the

variable RFI contamination levels between frequency sub-bands.

Incoherent beamforming has been demonstrated, and shown to afford less than the factor $\sqrt{N}$ sensitivity increase (where N is the number of combined telescopes) implied by the radiometer equation (e.g., Figure 6.11). This is unsurprising as e-MERLIN's telescopes do not have equal SEFDs (see Table 4.1), and due to the variable RFI contamination levels between telescopes.

The first successful observations of single pulses for RRAT J1819−1458 have been presented (see Figure 6.14), demonstrating the potential of LOFT-e's fast transient single pulse search pipeline. Due to technical issues only two telescopes successfully provided data to LOFT-e's computing nodes during these observations. Additionally, the phase calibrator selected was too far from the source to provide phase coherence. As such no imaging of the source was attempted. Section 6.2.2 presents plans for a future observation to demonstrate localisation of this RRAT by combining LOFT-e's single pulse search pipeline with e-MERLIN's 1 s imaging capability.

The sensitivity losses identified during LOFT-e observations show that in its current state, LOFT-e will be unlikely to detect FRBs if operating commensally. Therefore investigation into the nature of these losses will be an important contribution to future development of LOFT-e. Section 6.3 presents early investigation into the improvement to Gaussianity of data distributions which is afforded by LOFT-e's current AGC implementation technique, and into the overall effectiveness of the AGC. The AGC is shown to improve the Gaussianity of data distributions, however distributions do not consistently reach ideal levels. Additionally, an inconsistent response to AGC triggering has been identified for different e-MERLIN telescopes, sub-bands and polarisations. This includes systematic offsets in the data distributions measured between polarisations for the Defford and Knockin telescopes at 1414.4 MHz (e.g., Figure 6.18).

Finally, early investigation into the stability of data distributions post AGC-triggering has taken place. Data distributions are observed to be relatively stable for up to 155 minutes after observations. Statistical tests will be necessary to quantify these results.

8.4   On the DM-redshift analysis of FRBs

While the number of FRBs which will be localised to host galaxies will increase as new and upgraded telescopes with sub-arcsecond localisation capability become fully operational[3] the catalogue of unlocalised events will continue to grow. Chapter 7 presents

---

[3]Including, *one day*, e-MERLIN, hopefully!

work by the author aiming to maximise the usefulness of such unlocalised FRBs with poorly constrained redshifts using DM-redshift analysis.

In this chapter, physically motivated models for the electron distributions of spiral and elliptical galaxies are presented (e.g., Figure 7.1) and used to model probability distribution functions describing the relative likelihood that an FRB originating in such a galaxy will contribute a particular dispersion measure in the rest frame to the overall observed dispersion measure DM_{obs} of the source. This is achieved through modelling populations of galaxies with sources distributed according to physically motivated progenitor models. Such models include those tracing stellar populations (e.g., of OB stars and pulsars) of different ages, and galactocentric distributions. We conclude that host galaxies may contribute $>700\,\mathrm{pc\,cm^{-3}}$ units of dispersion measure in the rest frame depending on assumed progenitor models.

Using the redshift dependence of the host galaxy contribution to the overall dispersion measure, these probability distributions are scaled to different redshifts. Via a probabilistic approach, these distributions are combined with a physically motivated, redshift varying function describing the relative likelihood that a source at redshift z_s will contribute a particular dispersion measure due to the IGM, DM_{IGM}. The method establishes a framework through which the relative likelihood that any FRB at source redshift z_s will contribute an overall excess dispersion measure DM_{exc} to that afforded by the Milky Way, DM_{MW}. This framework may be applied for any progenitor or IGM model desired. The resulting probability distributions are shown to vary for different progenitor scenarios and imply that the host contribution DM_{host} may still be significant out to $z_s = 3$ in some cases (e.g., Figure 7.2).

Through application of Bayes' theorem, a method for the calculation of the likelihood that an FRB of given excess dispersion measure DM_{exc} will originate from a source redshift z_s is established. Thus a framework for constraining the redshifts of FRBs more rigorously than is standard practice is produced. This method is applied to unlocalised, catalogued FRBs using various progenitor scenarios. The results are presented in Appendix A.2.

Finally, techniques involving DM-redshift analysis which may increase the scientific yield of unlocalised FRBs are qualitatively identified. These include: FRB progenitor distinction, FRB host galaxy candidate rejection, Milky Way DM uncertainty mitigation, and disentanglement of the various components which make up DM_{obs}. It is proposed that such techniques be considered for general use during the study of FRBs. However it is noted that the most accurate uncertainties rely both on the use

of physically accurate models of host galaxies, and of the density fluctuations of the IGM. It will also rely on knowledge of FRB progenitors, an important open question in the research field.

Appendix A

A.1   A user guide to LOFT-e

The LOFT-e pipelines form the basis of a targeted high-time-resolution observing backend for the e-MERLIN interferometer, and will be adapted for full commensality in the future. The viability of the system has been demonstrated through targeted observations of pulsars and RRATs using time obtained during previous e-MERLIN proposal cycles. Examples of the fully processed output of such observations are shown in Chapter 6.

The pipeline may in future be rolled out for general purpose use. As several steps are necessary to ensure correct implementation of the pipeline and to optimise output data quality, instructions for a general user of the pipeline are provided below:

A.1.1   Using the LOFT-e data capture pipeline

1. **Choose observing configuration:** An observing schedule should be agreed with e-MERLIN schedulers[1] in accordance with the experimental aims of the observation. An observing band (i.e. L-band, C-band, K-band) should be chosen, alongside target sources and durations of observations. Phase calibrator observations are necessary if incoherent beamforming is to be used in conjunction with coherent beamforming or imaging. If imaging will be utilised, correlator integration times should be specified. Particular telescopes and sub-band centre frequencies to be streamed to LOFT-e must be chosen according to experimental aims and technical/RFI constraints at the observatory. Data distributions should be spread evenly across computing nodes where possible.

---

[1]schedulers@jb.man.ac.uk

2. **Cross-reference metadata:** After an observing configuration is agreed upon, it may be cross-referenced against the VDIF export and observing schedule metadata being received by LOFT-e nodes using the command line calls specified in Section 4.3.1[2].

3. **Source software packages:** The data capture pipeline is run from LOFT-e's head node, `transient`, and utilises the current version of the pulsar software package, `PSRSoft`[3]. This must be sourced before initiating data capture[4].

4. **Specify observation window:** This is done within the data capture pipeline's end-user control program[5]. If the pipeline is run within this observation window, all streamed VDIF data will be captured and further processing may take place offline.

5. **Specify lead telescope:** This is done within the data capture pipeline's end-user control program. The lead telescope is used to determine whether e-MERLIN is successfully tracking sources. If a partial array is being used (e.g. if certain telescopes are undergoing maintenance) not all telescopes may be on source. The telescope chosen as lead must take part in observations. The LOFT-e pipeline operates under the assumption that if the lead telescope is functional and tracking, then so are all other telescopes.

6. **Check data quality:** To ensure the best mapping of 3.5-bit to 2-bit data within the correlator's baseline boards for maximum S/N recovery, e-MERLIN's AGC (see Section 4.2.4) should be utilised. Due to investigations discussed in Chapter 6, current recommended practice is to manually activate the AGC for a short time via a command line call[6]. This command will generate scaling factors for each telescope, frequency sub-band, and polarisation, and apply them during requantisation. The AGC command may behave unreliably between individual data streams (see Section 6.3). It is therefore advisable to:

   (a) Run the AGC command.

   (b) Wait for a short (~ 10 s) period.

   (c) Begin data capture and record a short (< 1 min) amount of data.

---

[2]Currently, e-MERLIN may be operated in 'Observe mode', which allows the scheduler to directly specify to the correlator sources to observe and which configuration to use. This mode does not automatically provide observation schedule metadata and so should be avoided.

[3]`http://www.pulsarastronomy.net/wiki/Software/PSRSoft` (Accessed: 25/09/2018)

[4]PSRSoft may be sourced via the command line code:

`.   /packages/pulsarsoft/next_version/login/psr.bashrc`.

[5]Most recent version: `/home/cwalker/loft-e/lofte_receivers/eMControl_11cam.py`.

[6]`curl -data-urlencode input='correlator.vdifAgcSet()' emproc1:8080console`

      (d) Cease data capture and visually inspect the Gaussianity of recorded data[7].

      (e) If data distributions prove unsatisfactory, steps a) through d) should be repeated.

7. **Begin observing:** When these steps have been performed and incoming data streams are deemed of acceptable quality, the end-user data capture control program should be restarted for recording observations.

A.1.2 Masking LOFT-e data products

During processing of LOFT-e data reviewed in this work, a tool utilising the output DSPSR data products generated by LOFT-e's data processing pipeline to create masks for improving data quality has been created. This tool is generally used to create data stream dependent masks for the initial data file of specific observations. These masks are then applied to all further data files captured during the observation. The tool is used as follows:

1. Using the PSRCHIVE program `psrzap`, manually mask contaminated frequency channels of an output DSPSR archive file, producing a `.psh` command script (see van Straten et al. 2012).

2. Run the mask creation tool, specifying the directory[8] containing the data and command scripts. Mask files will be written to the data directory.

    More comprehensive analysis of the permanent RFI features of e-MERLIN data on a telescope, frequency sub-band, and baseline board basis is necessary before standard masks may be developed and implemented confidently into all future observations.

A.1.3 Troubleshooting LOFT-e hardware

Ongoing renovations and upgrades at JBO may occasionally cause power failures which disrupt LOFT-e's computing system and data/metadata connections. In the event of a power cut, diesel generators will start within a minute of the outage. When mains power returns, the generators are switched off and the observatory is reconnected. Machines at JBO, including LOFT-e nodes, are generally connected to uninteruppted

---

[7]Using, e.g., this tool provided by Eskil Varenius:
`/share/nas1/LOFT-e/software/e-MERLIN_data_check/vdif_check_data`.

   [8]`/share/nas1/LOFT-e/software/analysis_pipeline/psh_to_rfimask_lofte.py`

power supplies which provide power for tens of minutes to ensure a smooth transition during the process.

If power supplies fail and any machines contributing to the LOFT-e signal chain shut down, they must be restarted manually by staff members at JBO. Relevant machines have been described in the previous sections. All are shown in the comprehensive LOFT-e backend flow diagram in Figure 4.4.

The large size of transmitted VDIF data frames between the e-MERLIN correlator and LOFT-e requires an ethernet Jumbo maximum transmission unit (MTU) of 9000 bytes to be enabled on LOFT-e's ethernet0 socket. Power failure complications have previously been found to disable this configuration upon restarting LOFT-e's computing nodes. If no data are being received to a computing node following a power outage, the configuration of its ethernet0 MTU should be confirmed[9].

A.2   Redshift constraints for unlocalised FRBs

---

[9]This can be done using the command `ifconfig`.

Table A.1: Source redshift estimates z_s for FRBs in the FRB catalog (accessed on 04/04/2018). Columns (from left to right) contain: 1) source name, 2) DM-derived redshifts calculated using standard methods (Petroff et al. 2016; Ioka 2003), and redshifts calculated using excess electron models with 3) a host galaxy component ignored, and 4-7) the progenitor scenarios described in Fig. 7.2. All excess electron model redshift uncertainties are drawn from the 99% confidence intervals of their respective PDFs.

FRB	z_s (DM-derived)	z_s (IGM only)	z_s (YPSR)	z_s (H=0.2 kpc)	z_s (1 × R_e)	z_s (H=0.1 kpc)
FRB170827	0.116	$0.133^{+0.124}_{-0.130}$	$0.088^{+0.123}_{-0.085}$	$0.318^{+0.141}_{-0.315}$	$0.083^{+0.124}_{-0.080}$	$0.318^{+0.141}_{-0.315}$
FRB150807	0.191	$0.217^{+0.145}_{-0.215}$	$0.173^{+0.142}_{-0.170}$	$0.003^{+0.141}_{-0.000}$	$0.152^{+0.220}_{-0.140}$	$0.003^{+0.141}_{-0.000}$
FRB160410	0.184	$0.212^{+0.144}_{-0.210}$	$0.173^{+0.170}_{-0.150}$	$0.003^{+0.138}_{-0.000}$	$0.152^{+1.180}_{-0.136}$	$0.003^{+0.138}_{-0.000}$
FRB010724	0.275	$0.318^{+0.161}_{-0.315}$	$0.278^{+0.271}_{-0.200}$	$0.003^{+0.228}_{-0.000}$	$0.253^{+1.235}_{-0.226}$	$0.003^{+0.228}_{-0.000}$
FRB130628	0.348	$0.398^{+0.172}_{-0.395}$	$0.357^{+1.260}_{-0.233}$	$0.068^{+0.246}_{-0.065}$	$0.318^{+1.285}_{-0.286}$	$0.068^{+0.246}_{-0.065}$
FRB120127	0.435	$0.497^{+0.183}_{-0.495}$	$0.458^{+1.310}_{-0.267}$	$0.182^{+0.236}_{-0.180}$	$0.422^{+1.320}_{-0.386}$	$0.182^{+0.236}_{-0.180}$
FRB121102	0.307	$0.353^{+0.164}_{-0.350}$	$0.307^{+0.224}_{-0.232}$	$0.003^{+0.262}_{-0.000}$	$0.288^{+1.245}_{-0.259}$	$0.003^{+0.262}_{-0.000}$
FRB140514	0.440	$0.502^{+0.183}_{-0.500}$	$0.458^{+0.272}_{-0.288}$	$0.188^{+0.256}_{-0.184}$	$0.422^{+1.305}_{-0.385}$	$0.188^{+0.256}_{-0.184}$
FRB170107	0.479	$0.542^{+0.190}_{-0.540}$	$0.502^{+0.464}_{-0.269}$	$0.238^{+1.450}_{-0.215}$	$0.468^{+1.320}_{-0.426}$	$0.238^{+1.450}_{-0.215}$
FRB110523	0.483	$0.547^{+0.190}_{-0.545}$	$0.507^{+1.320}_{-0.269}$	$0.238^{+1.460}_{-0.214}$	$0.468^{+1.330}_{-0.425}$	$0.238^{+1.460}_{-0.214}$
FRB160608	0.370	$0.422^{+0.174}_{-0.420}$	$0.378^{+0.246}_{-0.266}$	$0.107^{+0.231}_{-0.105}$	$0.343^{+1.290}_{-0.311}$	$0.107^{+0.231}_{-0.105}$
FRB110626	0.563	$0.642^{+0.198}_{-0.640}$	$0.597^{+0.273}_{-0.311}$	$0.333^{+1.500}_{-0.300}$	$0.537^{+1.410}_{-0.493}$	$0.333^{+1.500}_{-0.300}$
FRB010621	0.185	$0.212^{+0.144}_{-0.210}$	$0.173^{+0.170}_{-0.150}$	$0.003^{+0.138}_{-0.000}$	$0.152^{+1.180}_{-0.136}$	$0.003^{+0.138}_{-0.000}$
FRB150418	0.490	$0.557^{+0.190}_{-0.555}$	$0.512^{+0.259}_{-0.309}$	$0.253^{+1.460}_{-0.229}$	$0.473^{+1.345}_{-0.430}$	$0.253^{+1.460}_{-0.229}$
FRB131104	0.590	$0.667^{+0.204}_{-0.665}$	$0.627^{+0.300}_{-0.292}$	$0.367^{+1.505}_{-0.332}$	$0.573^{+1.395}_{-0.527}$	$0.367^{+1.505}_{-0.332}$
FRB010125	0.567	$0.642^{+0.198}_{-0.640}$	$0.597^{+0.273}_{-0.311}$	$0.333^{+1.500}_{-0.300}$	$0.537^{+1.410}_{-0.493}$	$0.333^{+1.500}_{-0.300}$
FRB130729	0.692	$0.782^{+0.213}_{-0.780}$	$0.737^{+0.271}_{-0.328}$	$0.487^{+1.555}_{-0.439}$	$0.672^{+1.475}_{-0.622}$	$0.487^{+1.555}_{-0.439}$
FRB090625	0.723	$0.818^{+0.219}_{-0.815}$	$0.777^{+0.295}_{-0.300}$	$0.517^{+1.580}_{-0.465}$	$0.717^{+1.485}_{-0.665}$	$0.517^{+1.580}_{-0.465}$

FRB						
FRB110220	0.758	$0.852^{+0.225}_{-0.850}$	$0.812^{+0.273}_{-0.336}$	$0.562^{+1.590}_{-0.505}$	$0.742^{+1.510}_{-0.690}$	$0.562^{+1.590}_{-0.505}$
FRB130626	0.738	$0.837^{+0.219}_{-0.835}$	$0.792^{+0.264}_{-0.354}$	$0.537^{+1.585}_{-0.482}$	$0.717^{+1.510}_{-0.665}$	$0.537^{+1.585}_{-0.482}$
FRB151230	0.769	$0.867^{+0.226}_{-0.865}$	$0.828^{+0.280}_{-0.324}$	$0.578^{+1.595}_{-0.518}$	$0.747^{+1.525}_{-0.694}$	$0.578^{+1.595}_{-0.518}$
FRB110703	0.893	$1.008^{+0.241}_{-1.005}$	$0.967^{+0.275}_{-0.365}$	$0.717^{+1.665}_{-0.576}$	$0.867^{+1.610}_{-0.798}$	$0.717^{+1.665}_{-0.576}$
FRB150215	0.565	$0.642^{+0.198}_{-0.640}$	$0.597^{+0.273}_{-0.311}$	$0.333^{+1.500}_{-0.300}$	$0.537^{+1.410}_{-0.493}$	$0.333^{+1.500}_{-0.300}$
FRB170922	0.888	$1.002^{+0.241}_{-1.000}$	$0.962^{+0.274}_{-0.371}$	$0.717^{+1.660}_{-0.582}$	$0.867^{+1.595}_{-0.803}$	$0.717^{+1.660}_{-0.582}$
FRB160317	0.705	$0.797^{+0.218}_{-0.795}$	$0.762^{+0.388}_{-0.266}$	$0.507^{+1.560}_{-0.457}$	$0.672^{+1.490}_{-0.621}$	$0.507^{+1.560}_{-0.457}$
FRB171209	1.204	$1.363^{+0.276}_{-1.360}$	$1.323^{+0.289}_{-0.572}$	$1.093^{+1.750}_{-0.643}$	$1.162^{+1.730}_{-0.681}$	$1.093^{+1.750}_{-0.643}$
FRB150610	1.227	$1.393^{+0.279}_{-1.390}$	$1.357^{+0.315}_{-0.359}$	$1.133^{+1.745}_{-0.662}$	$1.162^{+1.765}_{-0.650}$	$1.133^{+1.745}_{-0.662}$
FRB121002	1.296	$1.472^{+0.285}_{-1.470}$	$1.432^{+0.301}_{-0.513}$	$1.212^{+1.760}_{-0.684}$	$1.242^{+1.785}_{-0.652}$	$1.212^{+1.760}_{-0.684}$
FRB151206	1.458	$1.667^{+0.304}_{-1.665}$	$1.633^{+0.329}_{-0.418}$	$1.412^{+1.805}_{-0.727}$	$1.442^{+1.840}_{-0.642}$	$1.412^{+1.805}_{-0.727}$
FRB160102	2.153	$2.578^{+0.417}_{-0.384}$	$2.553^{+0.543}_{-0.365}$	$2.353^{+1.965}_{-0.614}$	$2.373^{+2.005}_{-0.590}$	$2.353^{+1.965}_{-0.614}$

A.3 The variance in dispersion measures from the IGM

For our probability distribution function describing the likelihood that an FRB at given source redshift z_s will contribute a factor of $\mathrm{DM_{IGM}}$ to the overall observed dispersion measure $\mathrm{DM_{obs}}$, we assume a Gaussian:

$$P(\mathrm{DM_{IGM}}|\, z_{\mathrm{s}}) = \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{DM}}}\exp\left(-\frac{(\mathrm{DM}-\mathrm{DM_{IGM}}(z_{\mathrm{s}}))^2}{2\sigma_{\mathrm{DM}}^2}\right), \quad (A.1)$$

with mean DM:

$$\mathrm{DM_{IGM}}(z_{\mathrm{s}}) = \int_0^{\chi(z_{\mathrm{s}})} \mathrm{d}\chi\, \frac{n_{\mathrm{e}}(z)}{(1+z)^2}, \quad (A.2)$$

where the $n_{\mathrm{e}}(z)$ is the 3-dimensional electron number density; and with variance:

$$\sigma_{\mathrm{DM}}^2(\mathrm{DM}, z_{\mathrm{s}}) = \int_0^{\chi_{\mathrm{s}}} \mathrm{d}\chi(1+z_{\mathrm{s}})^2 \bar{n}_{\mathrm{e}}^2(0)\int\frac{\mathrm{d}^2 k_{\perp}}{(2\pi)^2}P_{\mathrm{e}}(k_{\perp}, z_{\mathrm{s}}), \quad (A.3)$$

following (McQuinn, 2014). In these equations, $\bar{n}_{\mathrm{e}}(0)$ is the mean electron density at redshift 0, χ is the comoving distance at redshift z which factors out the expansion of the Universe, $\mathrm{d}\chi = c\,\mathrm{d}z/H(z)$ where c is the speed of light and $H(z)$ is the Hubble factor at redshift z, the matter power spectrum at redshift z for wavenumber k is $P_{\mathrm{e}}(k, z)$, and $k_{\perp}$ is the perpendicular component (Dodelson, 2003).

Following Eqs. 3, 4, and Appendix (A1)–(A7) in Ma & Zhao (2014), electron density may be written:

$$n_{\mathrm{e}}(z) = \frac{\chi_{\mathrm{e}}\rho_{\mathrm{gas}}(z)}{\mu_{\mathrm{e}}m_{\mathrm{p}}}, \quad (A.4)$$

where $\Omega_{\mathrm{b}} = 0.048$ is the fraction of baryon density (Planck Collaboration et al., 2016b), m_{p} is the proton mass, $\mu_{\mathrm{e}} \simeq 1.14$ is the mean mass per electron, $\rho_{\mathrm{gas}}(z) = \rho_{\mathrm{gas}}(z = 0)(1+z)^3 = \Omega_{\mathrm{b}}\rho_{\mathrm{cr}}(1+z)^3$ is the gas density at redshift z, where $\rho_{\mathrm{cr}} = 1.879 h^2 \times 10^{-29}\,\mathrm{g\,cm^{-3}}$ is the critical density of the present Universe (Dodelson, 2003) and

$$\chi_{\mathrm{e}} = \frac{1 - Y_{\mathrm{p}}(1 - N_{\mathrm{He}}/4)}{1 - Y_{\mathrm{p}}/2}, \quad (A.5)$$

where χ_{e} is the number ratio between ionized and total electrons, $Y_{\mathrm{p}} \simeq 0.24$ is the helium mass fraction, and N_{He} is the number of ionized electrons per helium atom. The value of N_{He} can be between 0 and 2, so χ_{e} ranges from 0.86 to 1. With the change

of variables:

$$d\chi = \frac{-c\,dz}{H(z)},\tag{A.6}$$

Eq. A.2 may be written more explicitly as:

$$DM(z_s) = \left(\frac{\Omega_b \chi_e \rho_{cr}}{m_p \mu_e}\right) \int_0^{z_s} c\,dz\frac{(1+z)}{H(z)}.\tag{A.7}$$

Substituting the relevant quantities, we have:

$$\begin{aligned}
DM(z_s) &= 950.05\,[\mathrm{pc\,cm^{-3}}]\left(\frac{\Omega_b}{0.048}\right)\left(\frac{\chi_e}{1.0}\right)\left(\frac{h}{0.67}\right)\\
&\times \int_0^{z_s} dz\frac{(1+z)}{E(z)},
\end{aligned}\tag{A.8}$$

where $E(z) = \left[\Omega_m(1+z)^3 + \Omega_\Lambda\right]^{1/2}$ (Mo & White, 2002).

Likewise, the sightline to sightline variance in the mean dispersion measure following McQuinn (2014) is:

$$\begin{aligned}
\sigma^2(DM) &= \int_0^{\chi_s} d\chi(1+z)^2 \bar{n}_e^2(0) \int \frac{d^2 k_\perp}{(2\pi)^2} P_e(k_\perp, z)\\
&= \frac{c}{H_0}\left[\frac{\chi_e \Omega_b \rho_{cr}}{\mu_e m_p}\right]^2 \int_0^{z_s} dz\frac{(1+z)^2}{E(z)} \int dk\,k P_e(k,z),
\end{aligned}\tag{A.9}$$

where:

$$\bar{n}_e(0) = \frac{\chi_e \Omega_b \rho_{cr}}{\mu_e m_p},\tag{A.10}$$

as indicated by Eq. (A.4).

Bibliography

Abdo A. A., et al., 2013, ApJS, 208, 17

Adámek K., Armour W., 2016, preprint, (`arXiv:1611.09704`)

Alpar M. A., Cheng A. F., Ruderman M. A., Shaham J., 1982, Nature, 300, 728

Andreoni I., Cooke J., 2018, preprint, (`arXiv:1802.01100`)

Baade W., Zwicky F., 1934, Physical Review, 46, 76

Backer D. C., 1970, Nature, 228, 42

Bannister K. W., et al., 2017, ApJ, 841, L12

Barsdell B. R., Barnes D. G., Fluke C. J., 2010, MNRAS, 408, 1936

Barsdell B. R., Bailes M., Barnes D. G., Fluke C. J., 2012a, MNRAS, 422, 379

Barsdell B. R., Bailes M., Barnes D. G., Fluke C. J., 2012b, in Ballester P., Egret D.,
   Lorente N. P. F., eds, Astronomical Society of the Pacific Conference Series Vol.
   461, Astronomical Data Analysis Software and Systems XXI. p. 37

Bartel N., Morris D., Sieber W., Hankins T. H., 1982, ApJ, 258, 776

Bassa C. G., et al., 2016, MNRAS, 463, L36

Bassa C. G., Pleunis Z., Hessels J. W. T., 2017a, Astronomy and Computing, 18, 40

Bassa C. G., et al., 2017b, ApJ, 843, L8

Bates S. D., Lorimer D. R., Verbiest J. P. W., 2013, MNRAS, 431, 1352

Bayes M., Price M., 1763, Philosophical Transactions of the Royal Society of London
   Series I, 53, 370

Baym G., Hatsuda T., Kojo T., Powell P. D., Song Y., Takatsuka T., 2018, Reports on Progress in Physics, 81, 056902

Bhandari S., et al., 2018, MNRAS, 475, 1427

Bhat N. D. R., Cordes J. M., Camilo F., Nice D. J., Lorimer D. R., 2004, ApJ, 605, 759

Bhattacharyya B., et al., 2018, MNRAS, p. 880

Bowers F. K., Whyte D. A., Landecker T. L., Klingler R. J., 1973, IEEE Proceedings, 61, 1339

Boyle P. C., CHIME/Frb Collaboration 2018, The Astronomer's Telegram, 11901

Brecher K., Fesen R. A., Maran S. P., Brandt J. C., 1983, The Observatory, 103, 106

Brown B. A., Bregman J. N., 2001, ApJ, 547, 154

Burke B. F., Graham-Smith F., 2002, An Introduction to Radio Astronomy: Second Edition

Burke-Spolaor S., Bailes M., Ekers R., Macquart J.-P., Crawford Fronefield I., 2011, ApJ, 727, 18

Burke-Spolaor S., et al., 2016, ApJ, 826, 223

Caleb M., et al., 2017, MNRAS, 468, 3746

Cameron A. D., Barr E. D., Champion D. J., Kramer M., Zhu W. W., 2017, MNRAS, 468, 1994

Champion D. J., et al., 2016, MNRAS, 460, L30

Chandrasekhar S., 1931, ApJ, 74, 81

Chatterjee S., et al., 2017, Nature, 541, 58

Chippendale A. P., et al., 2015, in 2015 International Conference on Electromagnetics in Advanced Applications (ICEAA), p. 541-544. pp 541–544 (arXiv:1509.00544), doi:10.1109/ICEAA.2015.7297174

Clarke N., D'Addario L., Navarro R., Trinh J., 2014, Journal of Astronomical Instrumentation, 3, 1450004

Colegate T. M., Clarke N., 2011, PASA, 28, 299

Connor L., Lin H.-H., Masui K., Oppermann N., Pen U.-L., Peterson J. B., Roman A., Sievers J., 2016, MNRAS, 460, 1054

Cooper B. F. C., 1970, Australian Journal of Physics, 23, 521

Cordes J. M., 2002, in Stanimirovic S., Altschuler D., Goldsmith P., Salter C., eds, Astronomical Society of the Pacific Conference Series Vol. 278, Single-Dish Radio Astronomy: Techniques and Applications. pp 227–250

Cordes J. M., Lazio T. J. W., 2002, ArXiv Astrophysics e-prints,

Cordes J. M., McLaughlin M. A., 2003, ApJ, 596, 1142

Cordes J. M., Wasserman I., 2016, MNRAS, 457, 232

Cordes J. M., Lazio T. J. W., McLaughlin M. A., 2004a, New A Rev., 48, 1459

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

D'Amico N., et al., 2001, X-ray Astronomy: Stellar Endpoints, AGN, and the Diffuse X-ray Background, 599, 598

Deller A. T., Tingay S. J., Bailes M., West C., 2007, PASP, 119, 318

Deller A. T., et al., 2011, PASP, 123, 275

Dennison B., 2014, MNRAS, 443, L11

Dicke R. H., 1946, Review of Scientific Instruments, 17, 268

Dodelson S., 2003, Modern cosmology

D'Cruze M., 2018, PhD thesis, The University of Manchester

Falcke H., Gorham P., 2003, Astroparticle Physics, 19, 477

Falcke H., Rezzolla L., 2014, A&A, 562, A137

Fan X., et al., 2006, AJ, 132, 117

Farah W., et al., 2018, MNRAS, 478, 1209

Farrell W. M., Desch M. D., Zarka P., 1999, J. Geophys. Res., 104, 14025

Garrington S., Beswick R., 2016, Astronomy and Geophysics, 57, 3.28

Garrington S. T., et al., 2004, in Oschmann Jr. J. M., ed., Proc. SPIEVol. 5489, Ground-based Telescopes. pp 332–343, doi:10.1117/12.553235

Gold T., 1968, Nature, 218, 731

Goldreich P., Julian W. H., 1969, ApJ, 157, 869

Gómez G. C., Benjamin R. A., Cox D. P., 2001, AJ, 122, 908

Gordon D., Brisken W., Max-Moerbeck W., 2016, in Behrend D., Baver K. D., Armstrong K. L., eds, International VLBI Service for Geodesy and Astrometry 2016 General Meeting Proceedings: "New Horizons with VGOS", Eds. Dirk Behrend, Karen D. Baver, Kyla L. Armstrong, NASA/CP-2016-219016, p. 187-192. pp 187–192

Hagen J. B., Farley D. T., 1973, Radio Science, 8, 775

Han J. L., Manchester R. N., van Straten W., Demorest P., 2018, ApJS, 234, 11

Hankins T. H., Eilek J. A., Jones G., 2016, ApJ, 833, 47

Hassall T. E., et al., 2012, A&A, 543, A66

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

Hey J. S., 1946, Nature, 157, 47

Heywood I., et al., 2016, MNRAS, 457, 4160

Hobbs G., Lyne A., Kramer M., 2004, in Camilo F., Gaensler B. M., eds, IAU Symposium Vol. 218, Young Neutron Stars and Their Environments. p. 431 (arXiv:astro-ph/0310093)

Hobbs G., et al., 2010, Classical and Quantum Gravity, 27, 084013

Hu H. D., Esamdin A., Yuan J. P., Liu Z. Y., Xu R. X., Li J., Tao G. C., Wang N., 2011, A&A, 530, A67

Hulse R. A., Taylor J. H., 1975, ApJ, 195, L51

Ioka K., 2003, ApJ, 598, L79

Jansky K. G., 1933, Popular Astronomy, 41, 548

Johnston S., van Straten W., Kramer M., Bailes M., 2001, ApJ, 549, L101

Kalberla P. M. W., Dedes L., 2008, A&A, 487, 951

Kaplan D. L., 2008, in Yuan Y.-F., Li X.-D., Lai D., eds, American Institute of Physics Conference Series Vol. 968, Astrophysics of Compact Objects. pp 129–136 (arXiv:0801.1143), doi:10.1063/1.2840384

Kaspi V. M., Beloborodov A. M., 2017, ARA&A, 55, 261

Katz J. I., 2014, Phys. Rev. D, 89, 103009

Keane E. F., 2016, MNRAS, 459, 1360

Keane E. F., Kramer M., 2008, MNRAS, 391, 2009

Keane E. F., Kramer M., Lyne A. G., Stappers B. W., McLaughlin M. A., 2011, MN-RAS, 415, 3065

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

Keane E., et al., 2014, SUPERB - A SUrvey for Pulsars and Extragalactic Radio Bursts, ATNF Proposal

Keane E. F., et al., 2016, Nature, 530, 453

Keane E. F., et al., 2018, MNRAS, 473, 116

Keimpema A., et al., 2015, Experimental Astronomy, 39, 259

Keith M. J., et al., 2010, MNRAS, 409, 619

Komesaroff M. M., 1970, Nature, 225, 612

Kramer M., 2016, International Journal of Modern Physics D, 25, 1630029

Kronberg P. P., Bernet M. L., Miniati F., Lilly S. J., Short M. B., Higdon D. M., 2008, ApJ, 676, 70

Laplace P., 1812, Théorie Analytique de Probabilités. Courcier, Paris

Lattimer J. M., Prakash M., 2001, ApJ, 550, 426

Law C. J., Jones G., Backer D. C., Barott W. C., Bower G. C., Gutierrez-Kraybill C.,
   Williams P. K. G., Werthimer D., 2011, ApJ, 742, 12

Law C. J., et al., 2015, ApJ, 807, 16

Lazio J. W., et al., 2014, PASP, 126, 196

Li D., Pan Z., 2016, Radio Science, 51, 1060

Loeb A., Shvartzvald Y., Maoz D., 2014, MNRAS, 439, L46

Lorimer D. R., 2011, SIGPROC: Pulsar Signal Processing Programs, Astrophysics
   Source Code Library (ascl:1107.016)

Lorimer D. R., 2013, in van Leeuwen J., ed., IAU Symposium Vol. 291, Neu-
   tron Stars and Pulsars: Challenges and Opportunities after 80 years. pp 237–242
   (arXiv:1210.2746), doi:10.1017/S1743921312023769

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

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

Lovell B., 1964, Nature, 203, 11

Lovell J. E. J., Jauncey D. L., Bignall H. E., Kedziora-Chudczer L., Macquart J. P.,
   Rickett B. J., Tzioumis A. K., 2003, AJ, 126, 1699

Lyne A. G., Smith F. G., 1968, Nature, 218, 124

Lyne A. G., et al., 1998, MNRAS, 295, 743

Ma Y.-Z., Zhao G.-B., 2014, Physics Letters B, 735, 402

Maan Y., van Leeuwen J., 2017, preprint, (arXiv:1709.06104)

Macquart J.-P., 2011, ApJ, 734, 20

Macquart J.-P., Johnston S., 2015, MNRAS, 451, 3278

Macquart J.-P., et al., 2010, PASA, 27, 272

Macquart J. P., et al., 2015, Advancing Astrophysics with the Square Kilometre Array (AASKA14), p. 55

Maíz-Apellániz J., 2001, AJ, 121, 2737

Mamon G. A., Łokas E. L., 2005, MNRAS, 363, 705

Manchester R. N., Taylor J. H., 1972, Astrophys. Lett., 10, 67

Manchester R. N., Hobbs G. B., Teoh A., Hobbs M., 2005, AJ, 129, 1993

Manchester R. N., Fan G., Lyne A. G., Kaspi V. M., Crawford F., 2006, ApJ, 649, 235

Marcote B., et al., 2017, ApJ, 834, L8

Marrone D. P., Moran J. M., Zhao J.-H., Rao R., 2007, ApJ, 654, L57

Masui K., et al., 2015, Nature, 528, 523

McCool R., et al., 2006, Radio Science Bull., Vol. 317 (June 2006) p. 9-18, 317, 9

McCool R., Bentley M., Argo M., Spencer R., Garrington S., 2008,] 10.1109/E-COC.2008.4729488

McLaughlin M. A., et al., 2006, Nature, 439, 817

McQuinn M., 2014, ApJ, 780, L33

Michilli D., et al., 2018, Nature, 553, 182

Mo H. J., White S. D. M., 2002, MNRAS, 336, 112

Mooley K. P., et al., 2016, ApJ, 818, 105

Mori K., et al., 2013, ApJ, 770, L23

Morris D. J., et al., 2002, MNRAS, 335, 275

Mottez F., Zarka P., 2014, A&A, 569, A86

Murphy T., et al., 2013, PASA, 30, e006

Nagase F., 1989, PASJ, 41, 1

Nan R., et al., 2011, International Journal of Modern Physics D, 20, 989

Narayan R., 1992, Philosophical Transactions of the Royal Society of London Series A, 341, 151

Neiman M. S., 1943, Proceedings of the IRE, 31, 666

Ng C., et al., 2017, preprint, (arXiv:1702.04728)

Nicastro F., et al., 2018, Nature, 558, 406

Obrocka M., Stappers B., Wilkinson P., 2015, A&A, 579, A69

Olausen S. A., Kaspi V. M., 2014, ApJS, 212, 6

Oppenheimer J. R., Volkoff G. M., 1939, Physical Review, 55, 374

Özel F., Freire P., 2016, ARA&A, 54, 401

Pacini F., 1967, Nature, 216, 567

Palaniswamy D., Li Y., Zhang B., 2018, ApJ, 854, L12

Pen U.-L., Connor L., 2015, ApJ, 807, 179

Petroff E., et al., 2015a, MNRAS, 447, 246

Petroff E., et al., 2015b, MNRAS, 451, 3933

Petroff E., et al., 2015c, MNRAS, 454, 457

Petroff E., et al., 2016, PASA, 33, e045

Petroff E., et al., 2017, preprint, (arXiv:1710.08155)

Piro A. L., 2016, ApJ, 824, L32

Planck Collaboration et al., 2016a, A&A, 594, A9

Planck Collaboration et al., 2016b, A&A, 594, A13

Ray P. S., Wood K. S., Wolff M. T., 2017, preprint, (arXiv:1711.08507)

Reifenstein E. C., Brundage W. D., Staelin D. H., 1969, Physical Review Letters, 22, 311

Romero G. E., del Valle M. V., Vieyro F. L., 2016, Phys. Rev. D, 93, 023001

Rubio-Herrera E., Stappers B. W., Hessels J. W. T., Braun R., 2013, MNRAS, 428, 2857

Schnitzeler D. H. F. M., 2012, MNRAS, 427, 664

Scoville N., et al., 2007, ApJS, 172, 1

Shao L., Dai Z.-G., Fan Y.-Z., Zhang F.-W., Jin Z.-P., Wei D.-M., 2011, ApJ, 738, 19

Spencer R. E., 2009, in 8th International e-VLBI Workshop. p. 29

Spitler L. G., et al., 2014, ApJ, 790, 101

Spitler L. G., et al., 2016, Nature, 531, 202

Stappers B., 2016, in Proceedings of MeerKAT Science: On the Pathway to the SKA. 25-27 May, 2016 Stellenbosch, South Africa (MeerKAT2016). Online at <A href="href=">href="https://pos.sissa.it/cgi-bin/reader/conf.cgi?confid=277</A>, id.10. p. 10

Staveley-Smith L., et al., 1996, PASA, 13, 243

Sun X. H., Han J. L., 2004, MNRAS, 350, 232

Tanimura H., Hinshaw G., McCarthy I. G., Van Waerbeke L., Ma Y.-Z., Mead A., Hojjati A., Tröster T., 2017, preprint, (`arXiv:1709.05024`)

Tauris T. M., et al., 1994, ApJ, 428, L53

Tendulkar S. P., et al., 2017, ApJ, 834, L7

The CHIME/FRB Collaboration et al., 2018, preprint, (`arXiv:1803.11235`)

Thomasson P., 1986, QJRAS, 27, 413

Thompson A. R., Emerson D. T., Schwab F. R., 2007, Radio Science, 42, RS3022

Thornton D., et al., 2013, Science, 341, 53

Totani T., 2013, PASJ, 65, L12

Vieyro F. L., Romero G. E., Bosch-Ramon V., Marcote B., del Valle M. V., 2017, A&A, 602, A64

Walker C. R. H., Ma Y.-Z., Breton R. P., 2018a, preprint, (`arXiv:1804.01548`)

Walker C. R. H., et al., 2018b, preprint, (`arXiv:1804.01904`)

Wang H. G., et al., 2014, ApJ, 789, 73

Wayth R. B., Brisken W. F., Deller A. T., Majid W. A., Thompson D. R., Tingay S. J., Wagstaff K. L., 2011, ApJ, 735, 97

Weltevrede P., Johnston S., 2008, MNRAS, 387, 1755

Weltevrede P., Stappers B. W., Rankin J. M., Wright G. A. E., 2006, ApJ, 645, L149

Weltevrede P., Johnston S., Espinoza C. M., 2011, MNRAS, 411, 1917

Wolszczan A., Frail D. A., 1992, Nature, 355, 145

Woosley S. E., Heger A., Weaver T. A., 2002, Reviews of Modern Physics, 74, 1015

Xu J., Han J. L., 2015, Research in Astronomy and Astrophysics, 15, 1629

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

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

Yusifov I., Küçük I., 2004, A&A, 422, 545

Zavlin V. E., Pavlov G. G., 2004, ApJ, 616, 452

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

Zhou B., Li X., Wang T., Fan Y.-Z., Wei D.-M., 2014, Phys. Rev. D, 89, 107303

de Graaff A., Cai Y.-C., Heymans C., Peacock J. A., 2017, preprint, (`arXiv:1709.10378`)

van Straten W., Bailes M., 2011, PASA, 28, 1

van Straten W., Demorest P., Oslowski S., 2012, Astronomical Research and Technology, 9, 237