Observations of radio pulsars

and their physical implications

Henryk T. Haniewicz

School of Physics
University of East Anglia
July 2022

Thesis submitted in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy

Preface

Radio pulsars are fascinating objects of study. Pulsars are rapidly rotating, highly magnetised neutron stars which emit strong radio waves from their magnetic poles. At approximately ten kilometers in diameter, and with masses larger than our Sun's as predicted by general relativity, neutron stars necessarily comprise an ultra-compressed state of matter. At present, the conditions necessary to create this matter cannot be met in Earthican laboratories and current theories surrounding specific neutron star composition and evolution remain highly theoretical, through use of *equations of state*, with little reproducible empirical evidence. Radio pulsars are but one piece of the puzzle. Recently gravitational waves (GWs), amplified by merging compact binaries, have been detected. Analyses of these results and observations from other parts of the EM spectrum (e.g. from thermal sources) will help pin down viable equations of state so that we can better understand how our universe evolves.

Observed changes in pulsar pulse frequency, which is otherwise relatively stable, can sometimes be well modelled by assuming the neutron star as part of a binary system. Binary systems offer a plethora of information for the astronomer – with enough data, and under a given theory of gravity, the masses of the binary constituents can be measured which can lead to tighter constraints on that given gravitational framework. Measuring as many astrometric and orbital parameters as possible for each pulsar alongside using data from other binary sources (such as X-ray binaries), can constrain evolutionary models for the system post-first and post-second (if applicable) supernovae. Studies show that pulsars with higher spin frequency (more than about twenty unique pulses per second) are almost always found in binary systems. This is well explained by the theory of pulsar spin-up: matter is accreted to the neutron star from the companion which increases the neutron star's rotation speed due to the conservation of angular momentum – this process is called recycling.

If the pulsar in a binary has a spin period in the tens of milliseconds, the recycling process may not have been so rapid which may be implicit that the companion is also a neutron star. Direct companion searches are often fruitless due to the unlikely nature of seeing a radio pulsar signal, with only one previously observed companion pulsar. Double neutron star systems provide some of the most stringent tests of gravity, however their population is small and the key indicators for their evolution are poorly constrained. The following work presents and discusses the results from observations of five radio pulsar signals, all with rotation frequencies between fifteen and forty-four rotations per second.

PSR J1829+2456 is a member of one such double neutron star system in a compact binary orbit with a period of 1.18 days. Observations of this pulsar were made for three consecutive years, starting in 2017, following a hiatus in observations for almost ten years. Timing analysis of this new observing campaign shows a measureable signal delay due to the eclipsing of the neutron star by its companion, known as Shapiro delay, as well as a measureable two-dimensional shift in right ascension and declination – i.e. proper motion. In conjunction with previously known relativistic effects and assuming general relativity

as the correct theory of gravity, the masses are found to be 1.306 ± 0.007 Solar masses[1] and 1.299 ± 0.007 Solar masses for the neutron star and its companion, respectively. The proper motion implies a low-to-moderate space velocity of $\sim 50\,\mathrm{km\,s^{-1}}$ which, in tandem with the measured masses and mild orbital eccentricity, suggests the system is the result of a low-kick, symmetric supernova having undergone an ultra-stripped iron core-collapse.

PSRs J1851+0010 and J1853+0008, found by the PALFA collaboration, are also believed to be members of double neutron star systems although current timing analysis renders this claim inconclusive. Binary models suggest these two pulsars have rather wide orbital periods of 2.00 and 9.61 days, respectively as well as high eccentricities relative to the entire binary pulsar population. The relativistic effect of the binary constituents on their longitudes of periastron have been measured precisely and led to total mass measurements of 2.59 ± 0.02 and 2.6 ± 0.4 Solar masses respectively using general relativity, as well as precise component masses for PSR J1851+0010 whose companion mass is 1.15 ± 0.04 Solar masses. If the true companion mass is found to be close to its median value, this will be the lowest mass NS discovered to-date. At the present time, no significant proper motion measurement can be obtained so the evolution of these systems remains shrouded in some mystery.

PSR J1936+1805 is an isolated pulsar with an unusually fast rotation frequency (for an isolated pulsar) of about seventeen rotations per second. In lieu of any orbital fit, the rotation frequency is expected to be quite stable although the rate at which this frequency is changing, to a first order Taylor expansion, has been found to be $(-2.75 \pm 0.06) \times 10^{-17}\,\mathrm{rotations\,s^{-2}}$. These parameters are suggestive of binary pulsars, leading to the conclusion that this must be a disrupted binary – a pulsar that was once part of a high-mass binary system that did not survive either the supernova(e) nor the common envelope phases of evolution. The derived characteristic age of this pulsar is about ten-billion years meaning that it could have certainly been part of what would have otherwise become a double neutron star system.

PSR J1936+2142 is believed to be a member of a relatively circular ($e \sim 3 \times 10^{-5}$) close compact binary with an orbital period of 0.757 days. Phase connection of this pulsar shows it to have an observed pulse frequency decay of $(-4.6 \pm 0.6) \times 10^{-15}\,\mathrm{rotations\,s^{-2}}$ which implies the pulsar is old, with a characteristic age on the order of eleven billion years. Although some pulsars have been found with similar spin properties to these, they do not definitively categorise the binary nature of this system. Astrometric and orbital parameters, such as its very circular orbit suggest that this system is a high-mass neutron star-white dwarf binary, otherwise known as an intermediate mass binary pulsar as opposed to a double neutron star system, although a tighter orbital campaign will be necessary to conclude the evolution of this system.

The latter four pulsars represent the groundbreaking work being done by the Pulsar Arecibo L-band Feed Array (PALFA) survey collaboration. Its mission has brought about the discovery of many wonderful types of radio phenomena and shows the utmost value in collaborative efforts within astronomy. None of the present-day radio astronomy research

[1] All quoted uncertainties in this work are to the 2σ, or 95% confidence level, unless otherwise stated.

would be possible without global research efforts such as PALFA and other collaborations. I would personally like to thank the entirety of the PALFA collaboration for their ever thoughtful discussions and feedback concerning the PALFA discoveries. Of specific note, I would like to thank Emilie Parent for her unwavering help in helping to solve these pulsars.

More generally, I would like to thank Paulo Freire and Thomas Tauris for the multitude of thought provoking arguments about orbital mechanics and binary evolution over these past three and half years. This thesis would not be nearly as long nor interesting without your profound knowledge and input.

To my friends, who have been forced to endure 3+ years of what must have sounded like nonsense to the layman, I thank you dearly for your friendship. Indeed, there are too many of you to list here, but I hope this message and this thesis finds you all well.

I would like to thank my supervisor, Robert Ferdman, for his guidance and patience over these three years. Taking a chance on a student with qualifications in a different field must feel like a big gamble but I hope you feel, as much as I do, that it paid off. Thank you for every Sunday that you were available to talk at a moment's notice. Thank you for all the observations you saved by fixing my mistakes. Thank you for understanding and accepting my inexperience, allowing me to make my own mistakes whilst always offering calm and appropriate advice. Thank you for the opportunity to discover this field and all its wonders.

And finally, the person to whom I dedicate and owe this thesis, as well as my entire existence: my mother, Joanna Haniewicz. I have never known a life without love because I have always known you. Dziękuję.

Contents

List of Figures

List of Tables

List of Acronyms

AABBCRCI	Coherent PSRFITS format
AI	Artificial intelligence
AO	Arecibo Observatory
BCCE	Binary categorical crossentropy
BIPM	Bureau des Poids et Mesures
BT	Blandford-Teukolsky binary model
CCO	Central compact objects
CE	Common envelope
CHIME	Canadian Hydrogen Intensity Mapping Experiment
CNN	Convolutional neural network
DD	Damour-Deruelle binary model
DDGR	Damour-Deruelle GR binary model
DDH	Shapiro reparametrisation of DD
DM	Dispersion measure
DMAD	Double median absolute deviation
DNS	Double neutron star
DRP	Disrupted recycled pulsar
ELL1	Low eccentricity binary model
EM	Electromagnetic
FFT	Fast Fourier transform
GBT	Green Bank Telescope
GR	General relativity
GW	Gravitational wave
HMXB	High-mass X–ray binary
IF	Intermediate frequency
IMBP	Intermediate-mass binary pulsar
IMXB	Intermediate-mass X–ray binary
IQUV	Stokes PSRFITS format
ISM	Interstellar medium
LMBP	Low-mass binary pulsar

LMXB	Low-mass X–ray binary
LO	Local oscillator
ML	Machine learning
MS	Main sequence
MSP	Millisecond pulsar
NN	Neural network
NS	Neutron star
PALFA	Pulsar Arecibo L-band Feed Array
PSPM	Penn State Pulsar Machine
PSR	Pulsating source of radio
PUPPI	Puerto Rico Ultimate Pulsar Processing Instrument
QSO	Quasar – quasi-stellar object
RFI	Radio frequency interference
RLO	Roche lobe overflow
RMS	Root mean squared
RNN	Recurrent neural network
RRAT	Rotating radio transient
RVM	Rotating vector model
S/N	Signal-to-noise ratio
SGD	Steepest gradient descent
SSB	Solar system barycenter
TOA	Time-of-arrival
TT	Terrestrial time
VC	Vapnik–Chervonenkis (dimension)
WAPP	Wideband Arecibo Pulsar Processor

Chapter 1

Introduction

Fritz Zwicky and Walter Baade's 1934 article on the nature of supernovae (Baade and Zwicky, 1934a) led to the hypothesis of a very exotic type of *stellar remnant*. The theory stated that supernovae could result in very fast spinning, incredibly dense stellar remnants with incredibly small radii ($\sim 10\,\mathrm{km}$). Since neutrons can be gravitationally packed more efficiently than nuclei containing charged particles (Rutherford and Nuttal, 1913; Chadwick, 1932), it was theorised that these stellar remnants most likely consisted of ultra dense 'neutronic' matter and hence were dubbed neutron stars (NSs; Baade and Zwicky, 1934b). Current theory suggests that supernovae of progenitor stars of between 9 and $\sim 25\,M_\odot$[1] can become neutron stars post–supernova. Stars with initial mass between 8 and $10\,M_\odot$ likely create neutron stars through the core-collapse of an ONeMg core and stars with initial mass grater than $10\,M_\odot$ likely end their lives in Fe core-collapse supernovae (Heger et al., 2003). If the remnant left behind has a mass greater than the Tolman–Oppenheimer–Volkoff limit (about $2.3\,M_\odot$; Shibata et al., 2019), then further collapse ensues, leading to black hole formation. The small radii of NSs (see e.g. Özel and Freire, 2016) makes them very difficult to discover at visible wavelengths; only a handful are observable in the optical bands (Mignani, 2009). Gamma ray astronomy, such as the study of gamma ray bursts (GRBs) from merging NS binaries, has given us a wealth of information concerning neutron star composition (Tsang et al., 2012). However, arguably by far the most successful technique in aiding our understanding of neutron stars has been through radio astronomy.

1.1 Neutron stars as radio sources

The first radio pulsar (Pulsating Source of Radio – PSR) was discovered by Jocelyn Bell Burnell in 1967 at the Mullard Radio Astronomy Observatory, University of Cambridge, where a celestial object was observed in the Vulpecula constellation that displayed incredibly regular radio pulses separated by 1.337 s. This came to be known as CP1919, for "Cambridge Pulsar" (now known as PSR J1921+2153; Hewish et al., 1968). At first, no theory was aptly linked to support these findings, however when Bell Burnell discovered a similar set of regular pulses, each precisely 1.19 s apart, from a completely different region

[1] $M_\odot$ is the Solar mass $= 1.98847(7) \times 10^{30}$ kg

of space (PSR B1133+16; Hewish et al., 1968; Lyne and Rickett, 1968), the game was on to find a common source. The regularity and short periods of these radio pulses emanating from optically empty regions of space gave rise to the suggestion that these pulses could originate from neutron stars (Gold, 1968), a theory that was confirmed later that year by the discovery of two similar radio sources in the Vela (PSR J0835−4510, 89.33 ms; Large et al., 1968) and the Crab (PSR J0534+2200, 33.50 ms; Staelin and Reifenstein, 1968) nebulae, the latter of which is the remnant of a supernova that was visible to Earth in 1054.

The currently accepted model of a radio pulsar is that a conic emission of synchrotron radio waves originates from a region above the NS's magnetic poles (Pacini, 1968; Rankin, 1983) (see Figure 1.1). These magnetic poles, unlike Earth's and those of most main sequence (MS) stars, can have a large misalignment angle from the spin axis, with older members of the NS population having relatively small misalignment and younger members displaying larger angles, as expected from theory (Nowakowski, 1983). This misalignment implies the radio emission can potentially 'sweep' into our line of sight periodically as the NS rotates, a model coined the 'lighthouse model'. The catastrophic supernova environments that give rise to NSs (Baade and Zwicky, 1934c) could not possibly produce such an initially stable structure. A total dynamical energy of oscillation of about 10^{53} ergs must be dissipated from the NS before stability is reached leading to the theory that emission of electromagnetic (EM) radiation could be the cause (Hoyle et al., 1964). Therefore, every NS is theorised to emit a strong stream of charged particles along the open magnetic field lines at the polar caps (Sturrock, 1971). The term 'pulsar', then, is an entirely human–centric term: a pulsar is simply a NS that we can see[2]. However, it is currently unknown to what extent factors other than beaming direction (such as emission mechanism or age) affect our ability to see pulsars.

[2]Some pulsars, such as PSR J0737−3039B, have now stopped emitting towards our line of sight due to geodetic precession, however their status as pulsars remains. See Perera et al. (2010)

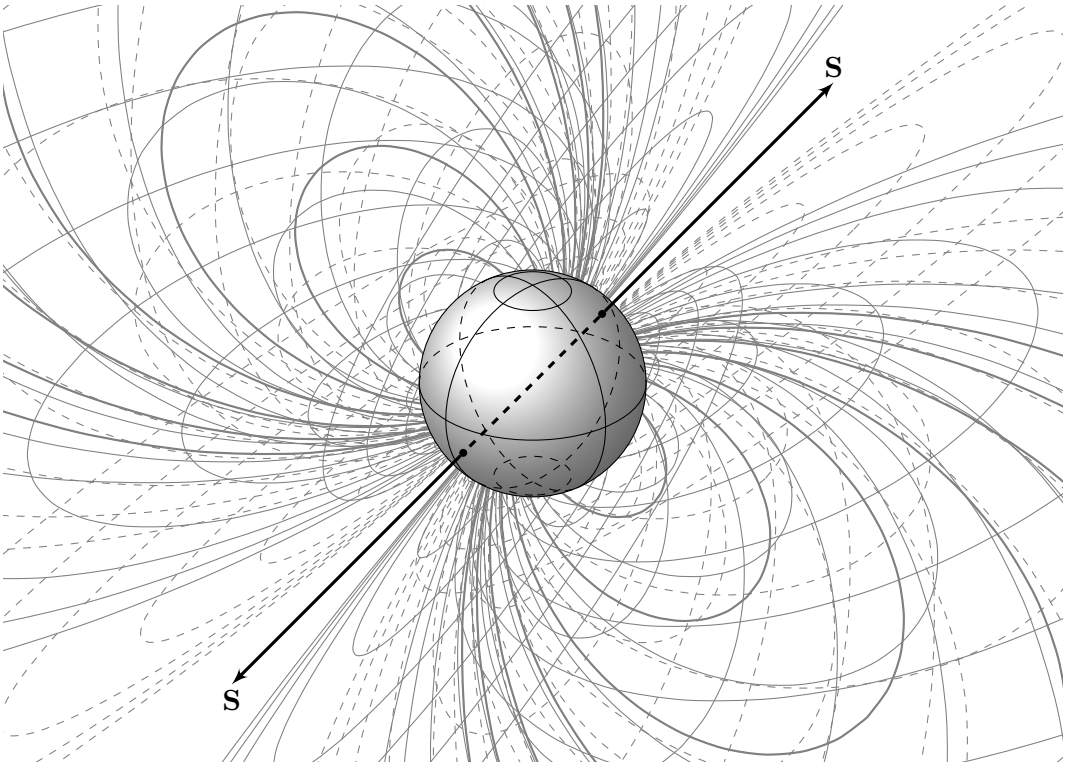


Figure 1.1: The basic schematic of a neutron star with electromagnetic jets being denoted with arrows labeled **S** to symbolise high flux. The cross points on the NS sphere represent the spin axis which is set here to 45deg from the magnetic axis as an example.

Since their discovery, over 2600 NSs have been discovered as pulsars (from any part of the EM spectrum) (Manchester et al., 2005), most of them within our own galaxy and a small percentage coming from globular clusters around the Milky Way. Notable current radio pulsar records are the fastest rotator (PSR J1748−2446ad, 1.396 ms; Hessels et al., 2006), slowest rotator (PSR J0250+5854, 23.5 s; Tan et al., 2018) and most massive (PSR J0740+6620, 2.14 $M_\odot$; Cromartie et al., 2019).

The so–called 'pulsar problem': the continued lack of definitive consensus on how pulsars are able to produce coherent radio emission (see e.g. Bahcall and Ostriker, 1997; Lorimer and Kramer, 2005), is perhaps one of limited data and relative newness of the field. As a scientific community, we have yet to come to a definitive consensus on even the most fundmental questions about how NSs might manifest as pulsars. Some questions that remain about NSs are: the origin, shape and mechanism of their radio emissions; their internal and atmospheric structure; and the nuances of additional pair creation due to the charge separation (Timokhin and Arons, 2012)[3]. It is therefore imperative that more pulsar studies can be carried out in order to probe these interesting phenomena. Added caveats are that pulsar spectral flux densities are incredibly weak, on the order of $1 - 10\,\mathrm{mJy}$.[4] – approximately the signal strength of an old mobile phone on the Moon as observed from Earth. This means that signals often compete with surrounding background interference. A final caveat is that pulsar signals are date-dependent (experiment results necessarily depend

[3]but cf. Philippov et al. (2020).
[4]1 Jansky (Jy) = $10^{-26}\,\mathrm{W\,m^{-2}\,Hz^{-1}}$

on the day on which the observations were taken). It is therefore important to understand the proposed signal theoretically before observation in order to disentangle the signal from the noise post-observation effectively, especially when the observation conditions are not ideal.

The rest of this section discusses the fundamental properties and theoretical understanding of pulsars that one must take into account when observations are taken.

1.1.1 Pulsar emission

Radio pulsar emission is made of highly elliptically polarised synchrotron radiation, meaning that it contains a combination of both linear and circular polarisation components, with the majority contribution coming from linear polarisation. Thermally radiating bodies that are still fusing low–mass elements in their cores (e.g. the Sun and other MS stars) emit, re-absorb and scatter their energetic photons through the process of radiation, usually modeled by a 'random walk' process (see e.g. Tunaley, 1974), in the core and by convection in outer layers of the star. Since there is no orientation bias, the MS star radiates in all directions essentially equally and the net polarisation of the emitted EM waves is ~ 0. As NSs are non-thermal radiators a degree of polarisation is expected, with the highest degree of linear polarisation observed for lower frequencies (Manchester, 1971). Observations at frequencies as high as 4.9 GHz have also confirmed the obverse situation to this: that the polarisation properties often diminish at high frequencies, leading towards depolarisation (Hoensbroech et al., 1998). Radio magnetars (see Section 1.1.4) are the exception to this and have been observed to be quite polarised even at high frequencies (Kramer et al., 2007).

Soon after the linking of radio pulsars to neutron stars, Radhakrishnan and Cooke (1969) suggested a model for the polarisation shifts observed in many pulsars (Manchester et al., 1975, is a notable example), in which the *parallactic position angle* of the linear polarised emission, with respect to the line of sight (Ψ), follows an 'S'-shaped curve as the beam sweeps into, and eventually out of, that line of sight, although this is not generally seen in recycled (see Section 1.1.3) systems (Stairs et al., 1999). The model is called the *rotating vector model* (RVM) and, although relatively weak, outputs from such a model have proven invaluable in the goal of helping to constrain orbital geometries for certain binary systems (see e.g. Konacki et al., 2003). The RVM can be summarised with a mathematical description of the parallactic angle using the fiducial plane as the rotational origin for phase, ϕ:

$$\tan(\Psi) = \frac{\sin(\alpha)\sin(\phi)}{\sin(\alpha+\beta)\cos(\alpha) - \cos(\alpha+\beta)\sin(\alpha)\cos(\phi)} \tag{1.1}$$

where α and β are the magnetic inclination and impact parameter respectively and, although possible to fit for directly, are most easily determined by the relation of steepest gradient on Ψ:

$$\left(\frac{d\Psi}{d\phi}\right)_{\max} = \frac{\sin\alpha}{\sin\beta}$$

The remainder of this section will go through the mathematical reasoning behind polarised feed setup. For a fixed direction of propagation, z, the electric vector $\mathbf{E}$, with magnitude $|\mathbf{E}|$, sweeps out an ellipse in the $x-y$ plane. This ellipse can be parametrised with respect to the Cartesian axis using the usual elliptical parameters for semi–major and semi–minor axes, A and B respectively, given as:

$$\begin{aligned} A &= \frac{|\mathbf{E}|}{\sqrt{2}}\sqrt{1+\sqrt{1-\sin^2(2\theta)\sin^2\beta}} \\ B &= \frac{|\mathbf{E}|}{\sqrt{2}}\sqrt{1-\sqrt{1-\sin^2(2\theta)\sin^2\beta}} \end{aligned} \tag{1.2}$$

where β is the signed difference in phase angle in the x and y directions ($\beta = \alpha_y - \alpha_x$ for Cartesian) and θ is the angle the semi–major axis makes with the x-axis. In this respect, $\beta = 0$ corresponds to complete linear polarisation ($A = |\mathbf{E}|$, $B = 0$) whereas $\beta = \pm 90°$ corresponds to circularly polarised light ($A = B = |\mathbf{E}|/\sqrt{2}$ if $\theta = 45°$). Figure 1.2 shows the geometry behind the polarisation.

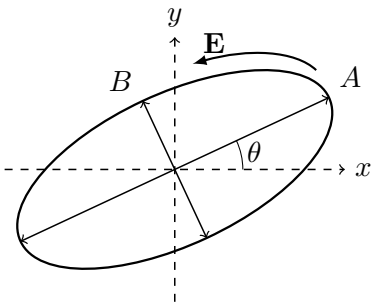


Figure 1.2: A diagrammatic representation of elliptically polarised light. A and B are the semi–major and semi–minor axes respectively. $\mathbf{E}$ is the EM field vector and the arrow, here arbitrarily orientated as RH, is the handedness of the circularly polarised light.

At an observatory, two orthogonal dipole feeds are typically installed in the receiver which are sensitive to the two components of linear polarisation. In order to obtain a circular dipole basis, the linear component receivers can be passed through a hybridisation procedure. The vectors from both dipole feeds are converted into a four-channel scalar basis which, for this thesis, is either coherent (uncalibrated) format (AABBCRCI) or Stokes calibrated format (IQUV) for folded data. In coherent format, each channel is a component of $A \otimes B$, with the vector description,

$$\left(A^2, B^2, \mathrm{CR}, \mathrm{CI}\right)$$

where CR and CI are the real co-spectrum and imaginary quadrature spectrum of A and B respectively (they represent the phase of polarisation). These four variables can uniquely determine a circular basis through the values of CR and CI. In Stokes format, the input channels are transformed into four geometrically motivated output channels representing the total light intensity, the degree of polarisation and the angular shape of the polarisation. These were designed by Stokes (1851) and were rediscovered almost 100 years later by

Perrin (1942) and Chandrasekhar (1947), of which the linear variant is given here:

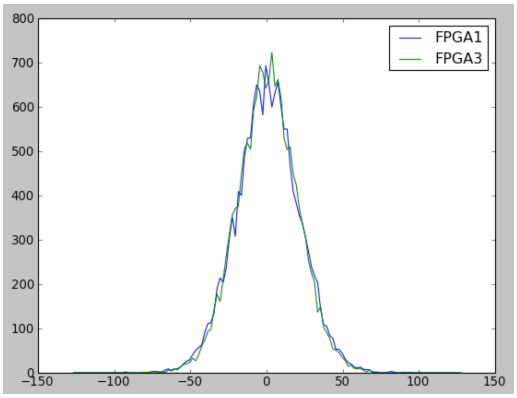
$$\begin{aligned} I &= A^2 + B^2 \\ Q &= \left(A^2 - B^2\right)\cos(2\theta) \\ U &= \left(A^2 - B^2\right)\sin(2\theta) \\ V &= 2ABh \end{aligned} \tag{1.3}$$

with θ defined as above, I is the total flux intensity and h is the so-called 'handedness' of the circular component which is $+1$ for right-handed polarisation and -1 for left-handed (International Astronomical Union, 1974). For the PSR J1829+2456 observations taken for this thesis, $h = +1$ and therefore $V = 2AB$. For the PALFA observations, the $h = -1$ handedness was used.

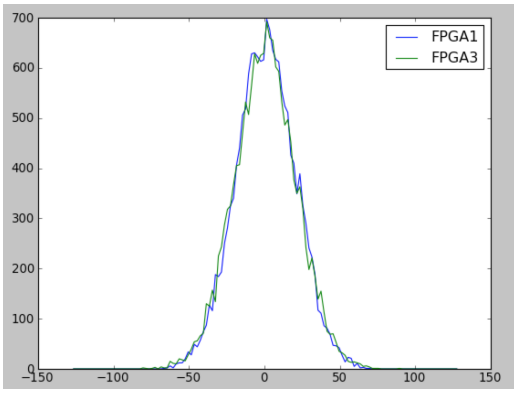
The data gathered and analysed for this work mostly uses the linear dipole feeds (although a few observations were taken with a circular basis), where the impulse response beam pattern $P(\phi, \lambda)$ measured in mJy, for aperture zenith angle ϕ and wavelength λ, agrees with the following one–dimensional aperture relation.

$$P\left(\phi, \lambda\right) = \frac{\lambda}{L}\left(\frac{\sin\left(\frac{\pi\phi L}{\lambda}\right)}{\pi\phi}\right)^2 \tag{1.4}$$

where L is the aperture size. A two–dimensional model sees the impulse response being determined by the antenna size in the current direction. One can see by Equation 1.4 that this profile ought to be roughly Gaussian (see Figure 1.3). Both polarisation channels for a given frequency *should* overlap on the beam pattern plot, as seen in Figure 1.3b but sometimes they do not, as seen in Figure 1.3a. When an observation takes place, the receiver beam pattern is entangled with signal from the source, so knowing by exactly how much this change in power has occurred, we can most accurately determine the contribution of signal due to the source.



(a) MJD 58615 (May 12$^{\text{th}}$, 2019)



(b) MJD 58845 (December 28$^{\text{th}}$, 2019)

Figure 1.3: Example power profiles, in mJy, from the 305-m Arecibo telescope's PUPPI backend, with respect to the zenith angle ϕ. The centre frequency for both observations was 430 MHz. The blue and green lines are the A(FPGA3) and B(FPGA1) feeds respectively. (a) The power levels for each dipole feed have a slight deviation at the apex. (b) The power levels are roughly the same for each feed.

1.1.2 Spin down

Although radio pulsars provide incredibly well defined pulse times owing to their stable spin periods (e.g. van Straten et al., 2001), their periods have been observed to increase over time, a process known as spin–down[5]. This is inferred by astronomers as a decrease in the pulsar's rotational energy, $E_{\rm rot}$, over time and is related to the spin period:

$$\frac{dE_{\rm rot}}{dt} = -\frac{1}{2}\frac{d(I\Omega^2)}{dt} = -I\Omega\dot{\Omega} = -4\pi^2 I\dot{P}P^{-3} \tag{1.5}$$

where Ω is the angular frequency of the pulsar ($\Omega = 2\pi/P$) and I is the NS moment of inertia which is typically assumed to be the canonical value of $\sim 10^{45}$ g cm^2 (Øvergård and Østgaard, 1991). The physical reason for this effect is assumed to be mostly from magnetic dipole emission and pulsar winds (Manchester et al., 1985; Blandford and Romani, 1988), although, admittedly, some radio emission must be accounted for even though this model assumes that the loss of radiation power (the radio emission we observe) over time is negligible when compared to $\dot{E}_{\rm rot}$, also known as the *spin–down luminosity*. Asuming that spin-down equations of motion take the general form of $\dot{\nu} \propto \nu^n$, the spin evolution due to magnetic dipole radiation for a pulsar can be modelled by the *braking index*, n:

$$n = 2 - \frac{P\ddot{P}}{\dot{P}^2} \tag{1.6}$$

where the magnetic dipole model predicts $n = 3$ (a derivation can be found in Appendix A.1). Most observations however show braking indices of much less than 3 (see e.g. Hamil et al., 2015), implying evolutionary deviations from the magnetic dipole model, and in practice it is unfeasible to precisely determine the braking index for many pulsars because only a handful have a measureable $\ddot{P}$. This is likely due to radio frequency interference (RFI) and noise contributions from the telescope being on the order of the proposed $\ddot{P}$ flux. This linear spin–down model, although crude, can give important extrapolative information about the pulsar's evolution such as its current age and magnetic field strength.

1.1.3 Characteristic age

From studies of spin–down and braking index, the *characteristic age*, $\tau_{\rm c}$, of the pulsar can be deduced. This is an approximation of the birth epoch of the NS (the time since supernova) by extrapolation of the current period assuming that the current period is much longer than the birth period and has always had a linear spin–down rate; this is given by (Jiang et al., 2013):

$$\tau_{\rm c} = \frac{1}{n-1}\frac{P}{\dot{P}} \tag{1.7}$$

These cannot be called true ages owing to the fact that their theoretical basis may not be strictly true; rather, it is a rough approximation and does not take into account any errors associated with the magnetic field and uncertainty as to the internal chemistry

[5]This is in contrast to the spin–up observed for NSs undergoing Case BB RLO (see Ritter and King, 2001, and Section 1.3)

and physics of a NS. The “true" age can be defined as the time period due to spin–down with respect to the birth period (Camilo et al., 1994, Equation (1)), but the required birth period knowledge is unknown. This measure can be resolved by taking into account the variation of magnetic moment over time (Helfand and Tademaru, 1977, Equation (3)). The associated error in the characteristic age estimate can be observationally determined through analysis of the Crab pulsar (Staelin and Reifenstein, 1968) which has a known age of 966 years and a derived characteristic age of ~ 1240 years, assuming the magnetic dipole model ($n = 3$) and an observed spin–down rate of 4.22×10^{-13}. This corresponds to a roughly 30% margin of error in τ_c, although young pulsars often show large discrepancies between their true and characteristic ages (Kaspi et al., 2001). Typical pulsar spin-down rates are between 10^{-15} and $10^{-19}\,\mathrm{s\,s^{-1}}$, giving characteristic ages in the millions and billions of years range and a much higher precision owing to their stability. Regardless of precision, this age estimate is still widely useful for pulsar population studies (see e.g. Dewey, 1992).

1.1.4 Surface magnetic field

Although braking index studies consistently show the magnetic dipole model of radiation to deviate significantly from reality, the surface magnetic field strength can be derived from Equation 1.5 assuming $n = 3$. Taking a classical electrodynamical approach, we can find the energy emission rate due to the dipole:

$$\frac{dE_{\mathrm{dip}}}{dt} = \frac{2}{3c^3}|\mathbf{m}|^2\Omega^4 \sin^2\alpha \tag{1.8}$$

where $\mathbf{m}$ is the magnetic field vector and α is the magnetic misallignment angle as defined in Section 1.1.1. The model assumes that $\dot{E}_{\mathrm{rot}} = \dot{E}_{\mathrm{dip}}$ so that Equations 1.5 and 1.8 can be equated, however this assumption is rather primitive as it does not take into account energy loss due to, for example, particle radiation or gravitional wave emission. Generally, pulsars with braking indices lower than $n = 3$ can be explained by particle emission (Ou et al., 2016) and a pulsar with $n = 5$ implies energy loss purely due to gravitational waves (see e.g. PSR J1640−4631, $n = 3.15 \pm 0.03$; de Araujo et al., 2016). Using the result of the dipole rotation energy loss assumption, and integrating Equation 1.6, one arrives at the angular frequency with a power law offset in n to account for observed deviations for the magnetic dipole emission model:

$$\dot{\Omega} = \left(\frac{2|\mathbf{m}|^2 \sin^2\alpha}{3c^3 I}\right)\Omega^3\alpha - \Omega^n \tag{1.9}$$

It can be shown that a reasonable approximation for the magnetic field moment is $B \approx |\mathbf{m}|/r^3$ (Jackson, 1998). Taking $r = R$ to be the NS star radius (approximately 10 km) and using the ideal magnetic dipole emission case and rearranging Equation 1.9, the surface magnetic field moment is found to be:

$$B_{\mathrm{surf}} = \sqrt{\frac{3c^3 I}{8\pi^2 R^6 \sin^2\alpha} P\dot{P}} \approx 3.2 \times 10^{19} G\sqrt{P\dot{P}} \tag{1.10}$$

although this is true necessarily only for when $\alpha = 90$deg and is a minimum; field strengths with α close to 0 are predicted to be a factor of 2 larger (Shapiro and Teukolsky, 1983). Given that the majority of pulsars display $P\dot{P} \sim 10^{-14}$ s^2 s^{-1}, this puts the expected $B_{\rm surf}$ in the 10^{12} G region.

It seems reasonable, then, to define three distinctive groups of NS pulsar based on the three afore mentioned principles; the *normal* pulsars, the *recycled* pulsars and the *magnetars*:

- Normal pulsars have observed spin periods ranging from about 100 ms – 4 s and tend to have high magnetic field strengths at their surface, on the order of 10^{12} G. Rotation and spin-down rates for normal pulsars mostly give $n = 3$ (magnetic dipole emission) characteristic ages between 100 Kyr and 1 Gyr.

- Recycled pulsars rotate in the the $1.5 - 100$ ms-period regime, have very stable rotation periods ($\sim 10^{-19}$ s s^{-1}) and lower surface magnetic field strength ($\sim 10^8$ G) than the normal pulsars, based on the magnetic dipole emission model. Pulsars with a very short rotation period ($\lesssim 30$ ms) are called *millisecond* pulsars (MSPs). Recycled pulsars generally have characteristic ages between 1 and 100 Gyr.

- Magnetars have long periods (> 1 s) and have the strongest estimated surface magnetic fields ($> 10^{14}$ G) based on the dipole emission model, however this model is largely inappropriate for current observations of magnetars which routinely imply braking indices < 3 (Lasky et al., 2017). The spin-down rates of megnetars are quite high ($\sim 10^{-12}$ s s^{-1}) compared with other pulsars, giving low characteristic ages of only a few thousand years, although spin-down rates are highly irregular, often best being modeled by glitch parameters or coupled with timing noise (see e.g. Turolla et al., 2015). One explanation for braking index measurements of magnetars is that they could have a twisted, toriodal magnetic field (possibly due to the merger of two NSs), which due to dynamical instabilities causes magnetic flux to 'bubble up' to the surface and back down again with an irregular node structure (i.e. not a dipole). Twisting increases the spin-down torque, implying smaller braking index (Thompson et al., 2002). Conversely, a twist-free theory of magnetar spin-down is given by Contopoulos and Spitkovsky (2006).

Further details concerning these groupings are documented by Soglasnov (2000). A convenient visual way to understand the different pulsar categories is by constructing a $P - \dot{P}$ diagram (e.g. Figure 1.4), which can be overlayed with magnetic field strength and age estimates. The distinction between the normal and recycled pulsars can be clearly seen in Figure 1.4 and the magnetars can often be identified from their slow periods. Many recycled pulsars have been observed in binary systems, leading to the theory that they have been "spun–up" ($\dot{P} < 0$) during a phase of accretion from the companion (Alpar et al., 1982). This is known as *recycling* and somehow contributes to the vast decrease in surface magnetic field strength observed for recycled pulsars when compared with normal

pulsars (Bisnovatyi-Kogan and Komberg, 1974); binary evolution is discussed in Sections 1.3 and 3.6 of this thesis. A fourth class of pulsar: the rotating radio transients (RRATs; McLaughlin et al., 2006) emit sporadically, often only once or twice per day, making them largely undetectable using Fourier domain searches. Through careful observation, it has been shown that RRATs are neutron stars (McLaughlin et al., 2006) and possibly magnetars.

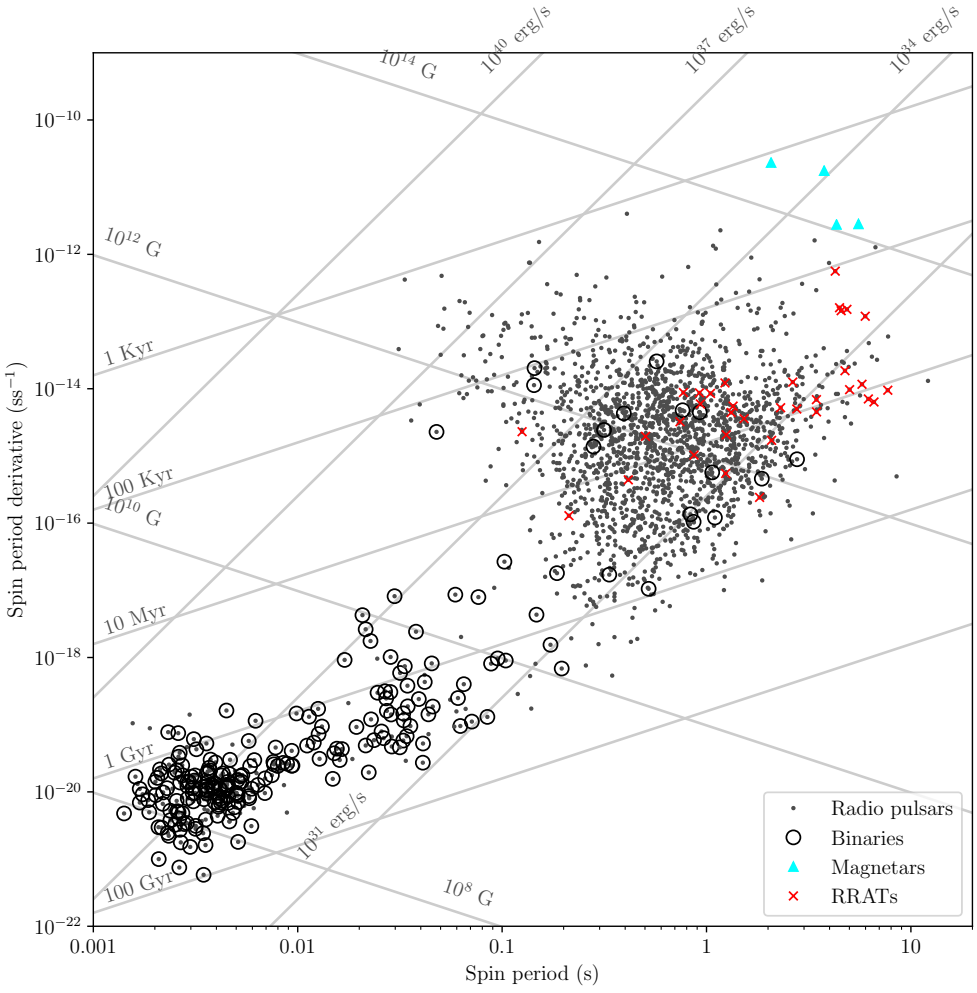


Figure 1.4: The $P - \dot{P}$ diagram for 2227 radio pulsars with data taken from the ATNF catalogue v1.64 (Manchester et al., 2005). Binary systems are highlighted by open circles, magnetars by cyan triangles and RRATs by red crosses. Only RRATs with phase-connected solutions have been included. Lines of constant magnetic field (in G), characteristic age (in yr) and spin–down energy loss rate (in erg/s) are overlayed.

1.2 Pulsar timing

Normal radio pulsars have observed spin–down rates in the $10^{-17} - 10^{-12}\,\mathrm{s\,s^{-1}}$ range. Essentially they can be viewed as stable rotators, but in comparison to MSPs which tend to display spin–down rates that are orders of magnitude smaller, the normal pulsars are generally unsuitable for ultra precise clock measurements, although normal pulsars are useful for population studies. Consequently, pulsar astronomers can rely on the period–stability of MSPs over a long epoch, implying that any delays with respect to the precise telescope clock are due to external effects. These timing delays can be the result of many phenomena such as solar system gravitational effects, the scattering and dispersive effects of the interstellar medium (ISM) and the relativistic orbital perturbations caused by binary systems (see e.g. Damour and Esposito-Farèse, 1992; Weisberg and Taylor, 2005, Chapter 3 of this work). It is the therefore the goal of pulsar timing to account for and therefore enable a precise investigation of a wide range of astrophysical phenomena with relatively few assumptions; this is the main concern of Chapters 3 and 4 in this work.

The pulse flux we observe due to the sweeping of the radio beam into our line of sight is referred to as the *pulse profile*. Most pulsars are not luminous enough to conduct accurate timing studies using pulses from each rotation. Adding to that, pulse shape from pulse-to-pulse is generally different but is stable over an long enough observation (see Figure 1.5) meaning that, over enough observations, conclusions can be drawn about the spin properties of the NS. The number of rotations needed for a stable profile ranges from a few hundred to a few thousand (Helfand et al., 1975) and is a function of the size, structure and emission angle of the radio beam. Stacking enough of the time series profiles modulo the pulse period and correct pulse phase should produce the true pulse shape, which is generally a function of the observer's line of sight, the observing frequency (discussed in Section 3.5) and the true structure of the emission beam. This process, called *folding*, produces an *integrated* pulse profile which is observed to be very stable for pulsars with low spin–down rate (Helfand et al., 1975). The high signal-to-noise ratio (S/N) of the integrated profile also results in a reduction in the uncertainty in the measured pulse time-of-arrival (TOA) that the profile represents. Many pulsar acquisition systems, or backends, offer real-time folding during observation but extra stacking may sometimes need to be done accompanied by RFI noise reduction. Pulses must be folded at a certain resolution (usually a power of 2) in phase, known as the sampling interval. Finer resolutions produce lower S/N profiles but contain more features whereas longer sampling intervals produce high S/N profiles post-folding.

Standard profile

In order to obtain accurate TOAs, one typically performs a cross-correlation between the observed profile, $p(t)$, and a very high S/N standard (or template) profile, $s(t)$. The standard profile is created by averaging many high S/N pulse profiles over an extended integration time, thus creating a profile averged over tens-of-thousands of rotations. In general, one ought to create a template that reflects the observing parameters, e.g. different

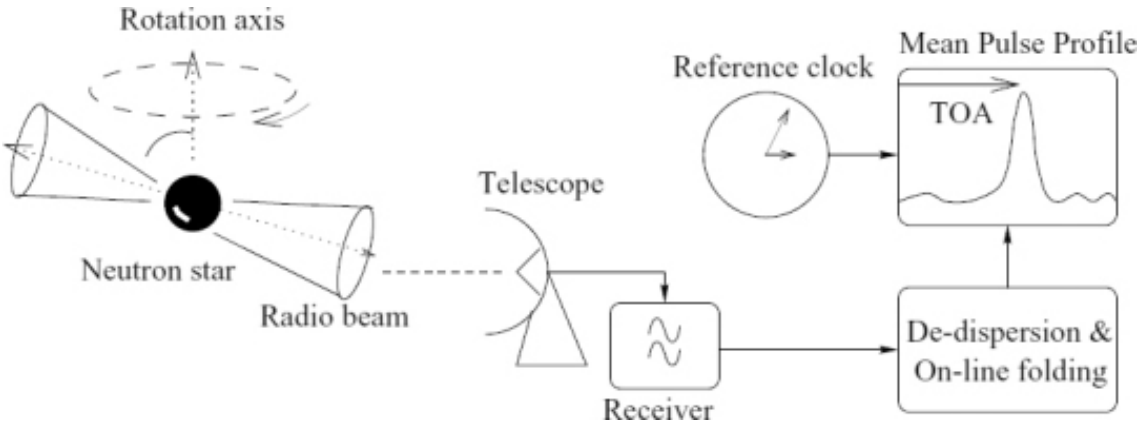


Figure 1.5: Schematic overview of a time-of-arrival procedure. Taken from Lorimer and Kramer (2005)©Reproduced with permission of Cambridge University Press through PLSclear.

templates for different frequency bands, telescopes, *et cetera* (Taylor, 1989). The goal of the template is to provide a fixed-phase profile against which to compare individual profile, $p(t)$, phases in order to measure their phase offset, τ, which can trivially be converted to a time offset by multiplying by the measured pulse period at the epoch corresponding to the profile. The observation will already have a time stamp associated with it, which is then added together with the time offset to get the pulse TOA. As more data is acquired for the pulsar, the standard profiles can be updated and TOAs from previous dates can be re-calculated given the updated standard profile. In this way, precise timing studies can assess the quality of updated templates and possibly suggest errors in analysis. The standard template model in the time domain is given by (Taylor, 1992):

$$p(t) = a + bs(t - \tau) + g(t) \tag{1.11}$$

where a is the possible baseline shift from the template to the pulse profile[6], b is a flux scaling factor so that variations in pulse flux are considered, and $g(t)$ encapsulates the contributions due to external noise in the time domain. The best match for the offset is found by minimising the output χ^2 value (Appendix A.2) of the trial phase offsets. In addition to the system noise, this method is unfortunately limited by the the sampling interval, which can often only provide accurate time offsets to the $\sim$ 10% level of the sampling interval (Taylor, 1992).

Pulsar timing is moving ever closer to nanosecond timing analyses with projects such as the ambitious International Pulsar Timing Array attempting to combine the data of 30 precisely timed pulsars in order to detect gravitational waves (Hobbs et al., 2010); and the Square Kilometer Array[7], which will aim to investigate extreme environments for general relativity (GR) such as exhibited by pulsars orbiting close to the Galactic centre (Smits et al., 2009). Because of this new era of technology, high precision timing, although

[6]In practice, $a \approx 0$ for properly baseline treated data.

[7]The SKA is currently not operational and scientific study on the instrument is not expected until 2027. Go to https://www.skatelescope.org/wp-content/uploads/2018/08/16231-Factsheets-operational-model-v4.pdf for recent news.

still feasible for most MSPs, can often not be accomplished as practically for normal pulsars using time-domain cross-correlation alone. In practice, all TOA calculations undergo frequency- (or Fourier-) domain cross-correlation as well (Taylor, 1992). This involves a χ^2 minimisation that utilises the discrete Fourier-transformed data profile and the template. The discrete Fourier transform of any pulse profile sampled with N_{BIN} intervals is given by:

$$\tilde{p}(\nu) = e^{-\mathrm{i}\theta_\nu} \sum_{k=0}^{N_{\text{BIN}}} p_k e^{2\pi \mathrm{i} k\nu / N_{\text{BIN}}} \tag{1.12}$$

with a similar result for $s(t) \rightarrow \tilde{s}(\nu)$. θ_ν is the particular phase of the pulse profile which will generally be different for $\tilde{s}(\nu)$, where the phase is denoted by ϕ_ν. TOAs created in this way are limited in accuracy by only the baseline noise of the data and template profiles and therefore enable better precision than time-domain cross correlations. Nuanced features of a pulse shape (such as sharp or multiple peaks) appear as higher harmonics in the frequency domain and can therefore be more easily taken into account. This is the reasoning behind the theory developed for RFI profile filtering.

Standard profiles have the further benefit of being more reliably modelled by a collection of pure functional forms. Widely used is a Gaussian summation (Foster et al., 1991; Kramer et al., 1994; Kramer, 1994) approach by which M Gaussian functions of generally different shapes and amplitudes are fit to the template (which has been created by other means). This has the added benefit of being able to cast a template into a different bin size (e.g. if there was any binning down during analysis). Mathematically, the template can now be represented by:

$$s(t) = \sum_{k=0}^{M} a_k e^{-\frac{t-\phi_k}{\sigma_k}} \tag{1.13}$$

where the k subscripted parameters (a, ϕ and σ) represent their usual Gaussian definitions of the amplitude, peak phase and the full width at half maximum, respectively, for the k^{th} Gaussian. This fitting method has been used for analysing templates in Chapters 3 and 4. The cyclic analogue to the Gaussian distribution function is the Von Mises distribution function (Mardia, 2000), which is related to the Bessel function of the first kind and is a good approximation to the wrapped Gaussian function. This can be used to create a template in the same way, with a more nuanced control sequence if the profile evolves significantly over observations (e.g. for mode-switching pulsars). Care must be taken with template building that involves the linear combination of several functions, so that one does not 'over-fit' the template; for example, small functional contributions often correspond to prominent interference features in the profile as opposed to true signal.

Fitting procedure

Arrival times calculated for many observation dates are collectively processed by applying them to a model that describes the rotation of the NS as a function of time. The resulting *timing residuals* denote how well the TOAs match the model that was fit to them. Residuals for the i^{th} TOA are first calculated before any fitting, known as *pre-fit* residuals:

$$R_i = \frac{\phi_i - \lfloor \phi_i \rceil}{\nu} \tag{1.14}$$

where ν is the spin frequency, $\lfloor \phi_i \rceil$ denotes the nearest integer[8] to the pulse phase, ϕ_i. The time evolution of the pulse phase, with respect to a reference time, t_0, can be written as a Taylor series with t taken to be the pulsar's proper time:

$$\phi(t) = \phi(t_0) + \sum_{k=1}^{\infty} \frac{\nu^{(k-1)}(t_0)}{k!}(t - t_0)^k \tag{1.15}$$

where $\nu^{(k)}(t_0)$ is the k^{th} derivative of the spin frequency at the reference epoch and the relationship between the spin frequency and the phase is $\nu = d\phi/dt$.[9] Given that it is unfeasible to measure values for $\ddot{\nu}$ for most pulsars except those undergoing rapid spin–down (see Ferdman et al., 2010, for a striking example of this), the phase evolution is most often (and usually necessarily) approximated by the first order term. The spin frequency and its derivatives are among the fittable parameters in this expansion, with the others arising from the measurement of t. The process of pulsar timing is to then measure and account for variations in the observed TOA, t_{obs}, from the expected arrival time at a particular spin phase. In the TEMPO2 pulsar timing package (Hobbs et al., 2006a; Edwards et al., 2006), which is used extensively in this work, the *post–fit* residuals are calculated by a least–squares algorithm, a lovely in-depth analysis of which is given by Coles et al. (2011). One can tailor TEMPO2 to perform non-linear fits but these do not significantly improve the timing solutions in this thesis. The least–squares approach aims to compartmentalise each pre-fit timing residual, R_i, into its deterministic component (the vector corresponding to each known parameter, P_j, together with a matrix describing the model, $\mathbf{M}$), and a random error component, which is called the *post-fit* residual, ϵ_i, given by:

$$R_i = M_{ij} P_j + \epsilon_i \tag{1.16}$$

Assuming the TOA uncertainty on the i^{th} pre-fit residual, σ_i, is roughly symmetrical about the $R = 0$ line, the fitting procedure can determine the most likely parameter values in $\mathbf{P}$ by minimizing the χ^2 value of the fit. The χ^2 does not take into account the degrees of freedom Γ allowed by the model, which is the number of TOAs minus the number of fit parameters; however, TEMPO2 does. The *reduced* χ^2 is given by the sum of the squared ratios of the pre-fit residuals to their errors, assuming all of the residuals are independent (not covariant with each other), and further dividing by Γ:

$$\chi^2_{\text{r}} = \frac{1}{\Gamma} \sum_{i=1}^{N_{\text{TOA}}} \left(\frac{R_i}{\sigma_i} \right)^2 \tag{1.17}$$

With these degrees of freedom taken into account, TEMPO2 uncertainties assume a χ^2_{r} at unity. A χ^2_{r} greater than 1 implies either an underestimate of parameter uncertainties in TEMPO2 which can be due to systematic TOA errors, or unaccounted physics promoting

[8] Hobbs et al. (2006a) use the notation N_i for the nearest integer. (cf. Hastad, 1988)

[9] To convert from telescope time to pulsar proper time, see Section 1.2.1 and subsequent sections.

the search for a better parameter fit. If a suitable minimum $\chi^2_{\rm r}$ can be found for the input data and the pre-fit residuals follow a Gaussian distribution, then the most likely set of physical parameter estimates, $\mathbf{P}_{\rm fit}$, can be determined via linear algebra manipulation (see e.g. Coles et al., 2011):

$$\mathbf{P}_{\rm fit} = (\mathbf{M}^{\rm T}\mathbf{M})^{-1}\mathbf{M}^{\rm T}\mathbf{R} \tag{1.18}$$

where T is the transpose matrix and $\mathbf{R}$ is the pre-fit timing residual vector. Given a fixed set of TOAs (implying a fixed $\mathbf{R}$ vector), the best parameter estimates are then simply a function of the model. Models that take into account more astrophysical phenomena, whilst still agreeing with theory, therefore produce the most accurate and precise parameter estimates. Ultimately this boils down to a two-pronged attack: the creation of valid TOAs and the use of a valid model. The pulsar proper time, t, can therefore be expressed as the sum of the observed TOAs and a series of 'delays' (collectively Δ) which can be implemented into the timing model:

$$t = t_{\rm obs} + \Delta \tag{1.19}$$

Delays that are not accounted for cause a significant and often periodic shift in the timing residuals and each delay parameter comes with a mathematical description to remove its effect, a process akin to baseline shifting. It is also in accounting for these delays that binary parameters may be measured. An outline of the various astrophysical delays accounted for in this work are outlined in the following sections.

1.2.1 Clock corrections

TOAs are recorded against local observatory clocks which vary significantly in their uniformity over long periods of time, making them unsuitable for pulsar timing as is. Clock corrections, $\Delta_{\rm clk}$, can be applied by the consistent monitoring of terrestrial time (TT) offsets between pairs of clocks. Clock corrections have the aim of transforming measurements into the Geocentric Celestial Reference System whose time co-ordinate (TCG) has the units of the SI definition of the second (Hobbs et al., 2006a, Section 3.1). Since the Earth is a rotating body adhering to effects due to relativity, terrestrial clocks do not run at the same rate as TCG. To overcome this, a realisation of TT containing a constant offset to take into account the value of the second is often used. The TT clock correction used for this thesis is the 2019 Bureau des Poids et Mesures, or TT(BIPM19), realisation (Guinot, 1988; Petit, 2004) which takes into account the daily dilation effect on TT and includes dates up until MJD[10] 58839. The origin point for TT(BIPM) is defined as the time offset of +32.184 s from the atomic clock standard, the Temps Atomique International (TAI), on January 1$^{\rm st}$ 1970[11]. The corrections are on the order of microseconds.

[10] Modified Julian Date, or MJD, counts the number of days since midnight on November 17$^{\rm th}$ 1858.
[11] ftp://ftp2.bipm.org/pub/tai/ttbipm/TTBIPM.2019

1.2.2 Solar system delays

The Solar System contains many bodies which affect our measurements based on the celestial positions, motions and gravitational effects of those bodies with respect to the topocentric (telescope) reference frame. The most prominent TOA delays caused by the Solar System (collectively $\Delta_{\odot}$) are (Lorimer and Kramer, 2005): the light–time variation due to the motion of the Earth around the Solar System barycentre (Rømer delay, $\Delta_{\mathrm{R}\odot}$); the delay due to signals passing through a large gravitational well such as that of the Sun or Jupiter (Shapiro delay, $\Delta_{\mathrm{S}\odot}$); and delays caused by the proximity of the Earth to other Solar System bodies manifesting in gravitational redshift and time dilation (Einstein delay, $\Delta_{\mathrm{E}\odot}$). Delays due to the Solar System are modelled using a Solar System ephemeris – a set of up-to-date parameters that describe the motions and masses within our Solar System and the Oort Cloud. The Solar System ephemerides used for timing in this thesis are the Developmental Ephemerides: DE435 (Folkner et al., 2016), which considers timing delays caused by Saturn; and DE438 (Folkner and Park, 2018), which provides updated orbits for Mercury, Mars, Jupiter and Cassini following from the completion of various planetary missions such as MESSENGER and JUNO (Solomon et al., 2007; Bolton, 2010). These ephemerides are both provided by the Jet Propulsion Laboratory. DE435 was used instead of DE438 for PSR J1829+2456 because DE438 was not out when the main analysis was run and comparison with DE438 outputs for this pulsar are identical. Compared with other sources of delay, the delays due to the Solar System carry the most weight when determining pulsar positions in the sky; the ephemeris must be able to take into consideration *at least* the effects in the current line of sight at the time of observation. Indeed, the location and proper motion (the *astrometric* parameters) of the pulsar can be determined by observing the deviation of the pulsar's position from the model when taking into account all of the above effects. The converse is also true, however; well timed and carefully analysed pulsars can be used to calibrate Solar System ephemerides by fitting to the known pulsar's parameters (Champion et al., 2010; Caballero et al., 2018).

1.2.3 Interstellar dispersion

Simultaneous pulsar observations taken at different recording frequencies have shown that emissions at higher frequencies arrive earlier than those at lower frequencies (e.g. Hankins and Rickett, 1986). This time delay, Δ_{D}, can be explained through the concept of *interstellar dispersion* of radio waves through the ISM. When EM waves propagate through the ISM, they have a group velocity, v_g, equal to the rate of change of angular frequency with respect to angular wavenumber; this is due to the slowing of the speed of light, c, in the interstellar plasma and is a function of the observing frequency (Manchester, 1977):

$$v_g(\nu) = c\sqrt{1 - \frac{q_e^2 n_e}{\pi m_e}\frac{1}{\nu^2}} \tag{1.20}$$

where $q_e \approx 1.6 \times 10^{-19}$ C and $m_e \approx 9.11 \times 10^{-31}$ kg are the charge and mass of the electron respectively, and n_e is the electron number density of the plasma – arguably the most difficult parameter to model here as it requires a robust Galactic model. The pre-

factor in front of the frequency term, $\nu_{\rm p}^2 = q_e^2 n_e (\pi m_e)^{-1}$, is known as the plasma frequency, and in the case of radio observing frequencies it is negligible (~ 2 kHz). If a magnetic dipole pulsar rotates at a frequency less than the plasma frequency then the refractive index of the ISM ($\mu = v_g/c$ for propagating radio waves) will be imaginary and therefore radio waves will not be able to propagate through the ISM. Using the approximation that $\nu_{\rm p} \sim 0$, the expected time delay for a wave propagating at frequency $\nu_{\rm low}$, from a distance d, with respect to a higher frequency, $\nu_{\rm high}$ is:

$$\begin{aligned}\Delta_{\rm D} &= \left(\int_0^d \frac{dl}{v_g(\nu_{\rm low})} - \frac{d}{c}\right) - \left(\int_0^d \frac{dl}{v_g(\nu_{\rm high})} - \frac{d}{c}\right) \\ &\approx \frac{q_e^2}{2\pi m_e c}\left(\frac{1}{\nu_{\rm low}^2} - \frac{1}{\nu_{\rm high}^2}\right)\int_0^d n_e dl \end{aligned} \tag{1.21}$$

where the vector l is the line of sight from Earth to the pulsar and the constant factor, stemming from the plasma frequency, is called the *dispersion constant* and has the approximate value 4.148808 ms (Lorimer and Kramer, 2005, Equation 4.6). Since the time delay is directly observable by analysing the raw time series over multiple frequencies, the sweep of the pulse peaks with frequency can be characterised by a single quantity known as the *dispersion measure*, DM. In the theoretically ideal case, the DM can be calculated by:

$$\mathrm{DM} = \frac{2\pi m_e c}{q_e^2}\frac{\Delta_{\rm D}}{\nu_{\rm low}^{-2} - \nu_{\rm high}^{-2}} \approx \int_0^d n_e dl \tag{1.22}$$

The implication of this is that pulsar distances may be estimated given a measured DM and a model for the Galactic free electron distribution (Cordes and Lazio, 2002; Yao et al., 2017). A reliable distance estimate can give workable estimations on the tangential velocities of pulsars (see Section 3.4.3) which can be used to fuel astrophysical arguments (e.g. system evolution based on supernova kick velocity). In reality though, DM is not a constant value due to regions of different electron densities in the ISM passing into and out of our line of sight over time (Kaspi et al., 1994a) or from the ionised envelope of a binary companion (Freire et al., 2003).

Precisely knowing the DM for a pulsar is vital to being able to properly fold the data so as to most accurately reconstruct the true pulse shape. Telescopes can only record data over a finite bandwidth and folded profiles that have not had the effects of DM accounted for will be broadened or 'smeared' at the trailing end of the pulse, leading to less precise TOAs. Correcting for the DM across a frequency band, known as *dedispersion*, has the effect of aligning the individual time-series pulses in phase, and is essential for high precision pulsar timing. The time delay, Δt, due to smearing across a band with bandwidth $\Delta\nu$ (in MHz) and centre frequency $\nu_{\rm ctr}$ (in GHz) is given by Cordes (2002):

$$\Delta t = \frac{8.3\mathrm{DM}\Delta\nu}{\nu_{\rm ctr}^3}\,\mu\mathrm{s} \tag{1.23}$$

As an example of the problematic nature of interstellar dispersion, PSR J1851+0010 ($P = 22.8$ ms; Chapter 4), with a DM of $107.7\,\mathrm{cm^{-3}\,pc}$, was observed on MJD 58130 with $\nu_{\rm ctr} =$

1.38 GHz and a bandwidth of 800 MHz containing 512 frequency channels. This amounts to a 272 ms delay over the whole band (although this equation does not generally hold for such large $\Delta\nu$) or a 0.53 ms delay over an individual channel – a detectable proportion of the spin period. Therefore, although further corrections to Equation 1.23 are needed for 800 MHz bandwidths, even the windows in which the equation is valid (over individual channels) can create unacceptable smearing.

Dedispersion can be carried out using either an incoherent (filterbank or autocorrelator) method as is seen in the search data in Chapter 4 or a coherent method, as can be seen in many recent timing studies (see e.g. Hankins, 2017, for a review). Incoherent dedispersion is the correction achieved by adding an artificial delay, calculated from Equation 1.23, to the signal in each frequency channel over the whole observing band. However, it does not account for dispersion within a given channel, and there will inevitably be some residual intra-channel smearing. In the above example of PSR J1851+0010, incoherent dedispersion can provide TOA precision of ~ 1 ms as a theoretical minimum, which is not ideal given its spin period. This cannot be used for timing when the expected pulsar feature sizes are on the order of this broadening.

Coherent dedispersion (Hankins and Rickett, 1975), such as is currently offered by most pulsar instruments at radio telescopes around the world, aims to resolve dispersion effects to the pulse shape within individual frequency channels in order to provide TOAs with better precision. Lorimer and Kramer (2005) provides an extensive account of coherent dedispersion but the main theory is as follows. The ISM is treated as a phase-only filter or *transfer function*, H. The (generally) complex voltage induced at the telescope receiver, $v_{\rm rcvr}(t)$, is the result of this phase shift applied to the propagating EM wave as emitted by the NS. To remove the dispersion, the receiver voltage is first Fourier transformed and then convolved with the inverse of the transfer function to give the Fourier transform of the voltage at the NS, $\tilde{v}_{\rm NS}(\nu)$ (in V μs). These Fourier transform integrals are finitiely bounded between $\pm\Delta\nu/2$ as the Nyquist sampling (Smith, 1997, Chapter 3) at the receiver ensures that signals outside this range are identically 0. The full equation is then:

$$\tilde{v}_{\rm NS}(\nu_{\rm ctr}+\nu) = \tilde{v}_{\rm rcvr}(\nu_{\rm ctr}+\nu)H^{-1}(\nu_{\rm ctr}+\nu) \tag{1.24}$$

The dedispersed voltage is then recovered by taking the inverse Fourier transform of Equation 1.24. The transfer function can be approximated (to the ns level) by:

$$H(\nu_{\rm ctr}+\nu) = \exp\left(\pm{\rm i}\frac{2\pi k_{\rm D}{\rm DM}}{(\nu_{\rm ctr}\pm\nu)\nu_{\rm ctr}^2}\nu^2\right) \tag{1.25}$$

where $k_{\rm D}$ is the dispersion constant. Although this method ideally removes all smearing, it is much more computationally expensive than filterbank dedispersion. Dispersion does also occur within the Solar System due to solar wind and correcting for this is handled by the Solar System emphemeris which takes the electron density to be constant in time and spherically symmetrical. Recent LOFAR studies suggest corrections to this model but these corrections do not affect pulsar timing results at the current level of precision (Tiburzi et al., 2021) so the electron density due to Solar wind in this work is assumed

constant in time with a value of $4\,\mathrm{cm}^{-3}$ at 1 AU, although Madison et al. (2019) have recently reported the Solar wind electron density to be $7.9 \pm 0.2\,\mathrm{cm}^{-3}$.

1.2.4 Orbital parameters

A binary orbit can be observed as a periodic change in delay in the arrival times, which are collectively denoted Δ_B (in much the same vein as the Solar System delays). Newtonian mechanics suggests that a binary orbit around a centre of mass can be described by five measureable orbital quantities known as the Keplerian parameters. These are:

- T_0, the epoch of periastron (sometimes the epoch of the ascending node), is the time stamp for the other orbital parameters. In much the same method as Equation 1.15, orbital parameters calculated at a previous known T_0 can be updated to reflect the current date.
- ω, the argument of periastron, is the angle between the ascending node of the binary system and the closest point of the orbit to its centre of mass (periastron). This is measured in the direction of motion and its reference plane is the orbital plane.
- e, the eccentricity, describes the ellipticity of the orbit as a dimensionless parameter. A bound system (i.e. a binary) will have values of e ranging from ~ 0 to < 1 with lower values for e describing more circular orbits.
- P_b, the orbital period, is the time taken (usually quoted in days) for one complete orbit of the system. This is generally calculated after long term periodic timing residual errors are detectable and resolvable as orbits may take a few days so good orbital coverage is necessary to accurately detect P_b.
- $x \equiv a_\mathrm{p} \sin i$, the projection of the orbit's semimajor axis a_p onto the plane of the sky as observed from Earth. Inclination angles are not resolvable by determination of x alone and the inclination must be calculated independently (see e.g. Equation 1.32).

These, along with the rate of change of the latter four parameters (discussed below) constitute the binary Rømer delay, Δ_RB as they serve to transform from the pulsar's reference frame to the centre of mass of the binary system (such as how $\Delta_{\mathrm{R}\odot}$ transforms from topocentric to barycentric).

These measurable quantities can be related to the system masses in a theory–independent way. A theory–independent timing model measures the timing delays as orbital parameters and corrections but does not infer masses from these measurements, due to their being no theory of gravity to relate masses to. To constrain masses from a model like this, one can measure multiple orbital parameters and calculate the mass function (Lorimer and Kramer, 2005), which is nominally a function of the pulsar and companion masses, m_p and m_c, as follows:

$$f(m_\mathrm{p}, m_\mathrm{c}) = \frac{4\pi^2}{T_\odot} \frac{x^3}{P_b^2} M_\odot = \frac{m_\mathrm{c}^3 \sin^3 i}{(m_\mathrm{p} + m_\mathrm{c})^2} \tag{1.26}$$

where $T_\odot \equiv G\,M_\odot/c^3 = 4.9254909476412675\,\mu$s is the conversion of the Solar mass into units of time. The mass function alone cannot be used to find the system component masses because it describes a ratio of masses, and m_p cannot generally be disentangled from that ratio. The component masses can be determined through observation of their effect on the Keplerian orbits and by employing a suitable theory of gravity (thus the model becomes theory–dependent). Additionally, the orbital inclination is generally unknown and must be found independently of at least one of the masses.

Post–Keplerian parameters

Most pulsar binaries fit the Keplerian orbital model well due to the relatively small orbital speeds predicted in these systems. In the case of two objects in a more compact binary system such as a double neutron star (DNS) or compact NS-WD, the approximate orbital speed is given by the *Vis–viva* equation:

$$v = \sqrt{GM_\mathrm{T}\left(\frac{2}{R} - \frac{1}{a_\mathrm{p}}\right)} \tag{1.27}$$

where R is the separation between the two bodies. Close DNS systems (such as e.g. Hulse and Taylor, 1975) have average orbital separations of only $\sim 2\,R_\odot$[12]. This gives them orbital speeds of $\sim 0.2\%$ of the speed of light at their closest approach which, in combination with the large masses present in compact binaries, result in measurable orbital relativistic effects on the observed pulse TOAs; these are designated *relativistic binaries*. These perturbations to the Keplerian orbits can be measured in a theory–independent way to give seven 'post-Keplerian' parameters (Damour and Deruelle, 1985; Damour and Deruelle, 1986):

- $\dot{\omega}$, the advance of periastron, is the secular change of the argument of periastron over time measured in deg yr^{-1}.
- γ, roughly equivalent to the Einstein delay Δ_EB, and the binary equivalent of $\Delta_{\mathrm{E}\odot}$. This parameterises the gravitational redshift and time dilation due to the varying separation of the pulsar from its companion.
- $\dot{P}_b$, the orbital decay, is the rate at which the orbital period decreases over time as the NSs move towards a merger event. This is caused by emission of quadrupolar gravitational waves (e.g. Abbott et al., 2017) predicted by relativity.
- r and s, the Shapiro "range" and "shape" parameters. r determines the extent of Shapiro delay due to the mass of the companion, as this determines the magnitude of the gravitational potential well through which the signal travels. s determines the extent of Shapiro delay with respect to the system's orientation and is equivalent to $\sin i$.
- δ_r and δ_θ parameterise the relativistic deformations of the orbit and cannot generally be measured separately.

[12] $R_\odot$ is the Solar radius: the radius at which the Sun's optical depth is at 2/3. $R_\odot = 6.957 \times 10^5$ km

Once measured in a theory–independent way, these PK parameters can be held at a fixed value in order to determine the system masses under specific theories of gravity. In the case of GR, the PK parameters can be parametrised in terms of the Keplerian orbital elements and the masses as follows (e.g. Edwards et al., 2006):

$$\dot{\omega}^{\rm GR} = 3T_{\odot}^{2/3}\left(\frac{P_b}{2\pi}\right)^{-5/3}\frac{(m_{\rm p}+m_{\rm c})^{2/3}}{1-e^2} \tag{1.28}$$

$$\gamma^{\rm GR} = T_{\odot}^{2/3}\left(\frac{P_b}{2\pi}\right)^{1/3} e\frac{m_{\rm c}(m_{\rm p}+2m_{\rm c})}{(m_{\rm p}+m_{\rm c})^{4/3}} \tag{1.29}$$

$$\dot{P}_b^{\rm GR} = -\frac{192\pi}{5}T_{\odot}^{5/3}\left(\frac{P_b}{2\pi}\right)^{-5/3}\frac{m_{\rm p}m_{\rm c}}{(m_{\rm p}+m_{\rm c})^{1/3}}f(e) \tag{1.30}$$

$$r^{\rm GR} = T_{\odot}m_{\rm c} \tag{1.31}$$

$$s^{\rm GR} \equiv \sin i = T_{\odot}^{-1/3}\left(\frac{P_b}{2\pi}\right)^{-2/3} x\frac{(m_{\rm p}+m_{\rm c})^{2/3}}{m_{\rm c}} \tag{1.32}$$

$$\delta_r^{\rm GR} = T_{\odot}^{2/3}\left(\frac{P_b}{2\pi}\right)^{-2/3}\frac{3m_{\rm p}^2+6m_{\rm p}m_{\rm c}+2m_{\rm c}^2}{(m_{\rm p}+m_{\rm c})^{4/3}} \tag{1.33}$$

$$\delta_\theta^{\rm GR} = T_{\odot}^{2/3}\left(\frac{P_b}{2\pi}\right)^{-2/3}\frac{\frac{7}{2}m_{\rm p}^2+6m_{\rm p}m_{\rm c}+2m_{\rm c}^2}{(m_{\rm p}+m_{\rm c})^{4/3}} \tag{1.34}$$

where $f(e)$ in Equation 1.30 is defined as:

$$f(e) \equiv \left(1+\frac{73}{24}e^2+\frac{37}{96}e^4\right)(1-e^2)^{-7/2} \tag{1.35}$$

Under GR then, it is possible to determine the total system mass from a significant measurement of one PK parameter and the individual component masses if a measurement of two or more are made. The PK-parameter for which the least data is needed is $\dot{\omega}$ and this parameter is often used to infer a GR-derived total mass (Equation 1.28). Depending on the orbital geometry, the second most easy to measure PK parameters are either Shapiro delay (for orbits with $\sin i \approx 1$) or Einstein delay. A measurement of $\dot{P}_b$ due to gravitational wave emission requires great orbital coverage and is therefore usually not detectable for the first few years of observation. Once the pulse properties are known, orbital coverage can be increased through planned dense campaigns in which the pulsar is observed more frequently than in a typical observing campaign (see Chapters 3 and 4).

Expanding on Equation 1.19, the total contribution to the delays as probed in this thesis is therefore:

$$t = t_{\rm obs} + \Delta_{\rm clk} - \Delta_{\rm D} + \Delta_{\odot} + \Delta_{\rm B} \tag{1.36}$$

This is the equation that defines the model used for timing analyses.

1.3 Binary systems containing pulsars

The observational study of NSs in binary systems began with the discovery of the binary pulsar PSR B1913+16 (PSR J1915+1606, 59.03 ms; Hulse and Taylor, 1975), also known as the Hulse–Taylor binary. Binary systems, which can sometimes be recognised as such directly from TOA residual deviations, can allow for precise tests of the predictions of GR in the strong-field regime (Taylor and Weisberg, 1989), as well as providing tests for alternative theories of gravity (e.g. Damour and Esposito-Farèse, 1992; Damour and Esposito-Farèse, 1993; Damour and Esposito-Farèse, 1996). Recently, the merger of close dense binaries containing NSs has been the topic of much scientific interest, owing to the first direct detection of a NS-NS merger from gravitational wave (GW) analysis (Abbott et al., 2017). GWs are ripples in spacetime caused by the close orbit of two very massive bodies. These waves solve the weak-field approximation of the Einstein equations in the vaccuum radiative regime, displaying linear propagation in this regime, but become highly non-linear in the strong-field regime around NSs and black holes, possibly violating GR in this regime (see e.g. Villatoro, 2018). GR suggests that orbits of compact bodies should shrink over time due to emission of GWs, a theory that is incredibly well constrained by the precise timing of the Hulse-Taylor binary – the impressive agreement with the $\dot{P}_b$ of PSR B1913+16 and the theoretical $\dot{P}_b$ due to GW emission is known simply as "the result". Extensive overviews of binary systems and NS binary evolution can be found in, e.g., Bhattacharya and van den Heuvel (1991), Tauris and van den Heuvel (2006), van den Heuvel (2007), and Tauris et al. (2017); a few fundamental concepts of binaries and their evolution are outlined below.

The life cycle of any NS binary system starts with two main sequence stars orbiting a common barycentre. The more massive of these stars, which is likely to be a massive but short-lived O or B-type star, terminates its main sequence in a core-collapse Type Ib/c supernova (Yoon et al., 2010), or possibly a Type II supernova. This leaves behind a NS, which may briefly be seen as a radio pulsar, orbiting a MS companion (as has been observed from e.g. PSR J0045−7319 (Kaspi et al., 1994b; Bell et al., 1995)), before being observable as an X-ray binary (XB). The type of XB formed after the first supernova is determined by the nature of the companion star and binary properties, which dictate the rate and volume of mass accretion, and are both strongly correlated with the final NS spin period, leading to the theory that current spin-orbit trends suggest categories of evolution groups (e.g. Tauris et al., 2012, and Figure 4.8 in this work). XBs are categorised by their companion masses, into high, intermediate and low-mass X-ray binaries (HMXB, IMXB, LMXB).

The XB stage sees the evolution of the MS companion which causes it to expand past its Roche Lobe if the Roche Lobe is smaller that the expa nsion limit (Roche-lobe overflow – RLO). To conserve angular momentum between the newly formed NS and the low-mass MS companion, a mass accretion process occurs from the MS companion to the NS via either a long-duration, stable mass transfer through the L1 Lagrangian point (in the case of most LMXBs), or via unconservative stellar wind (in the case of many HMXBs). This accretion process emits very strong X-ray signals as predicted via blackbody emission calculations.

Pulsars observed in binary systems do not have spin periods that would be expected for the majority of isolated pulsars[13] that have undergone spin-down since birth. Furthermore, most pulsars are thought to eventually reach spin periods so high ($P \sim 5$ – $10\,\mathrm{s}$ for a normal pulsar) that EM emission becomes unfavorable due to the diminished magnetic field being unable to align charged particles at the NS surface, and the pulsars would "turn off". What, then, is one possible explanation for pulsars which appear to have spin periods much shorter than their proposed birth period? The answer is *recycling* via mass accretion due to RLO of the companion star. This process can be divided into three main categories:

- *Case A* RLO occurs when the mass donor (companion) star is still burning hydrogen when RLO starts. Since LMXBs have low-mass ($< 8\,M_\odot$) companions that usually only burn hydrogen and helium (He), the majority of these binaries can only proceed via Case A RLO, eventually becoming a NS-He WD binary (see e.g. Podsiadlowski et al., 2002) and make up the majority of observed NS-WD binaries. The mass transfer in Case A RLO is a slow and conservative process through the L_1 Lagrangian point (the point between the two bodies where the gravitational influence of each are effectively cancelled) which results in efficient spin-up to millisecond periods. Hence, most of the MSPs most likely resulted from Case A RLO. Case A RLO can also give rise to NS-CO WD binaries if the progenitor system is a short ($\sim$ 1d) period IMXB with donor star initial masses of $3 - 5\,M_\odot$ (Tauris et al., 2000) and, in rare cases, IMXBs can form NS-He WDs if Case A RLO is initiated early enough.

- *Case B* RLO occurs if the RLO starts when the donor's core hydrogen burning phase is over. This happens when wider IMXBs of about $3 - 10$ days with donor masses $> 2.5\,M_\odot$ fill their Roche Lobe after evolving off the MS (Tauris et al., 2000). In the case where the donor star is $< 5\,M_\odot$, the resulting system is a NS-CO WD binary with orbital periods of about $3 - 50$ days (Tauris et al., 2000). IMXB donors exceeding $5\,M_\odot$ can undergo dynamically unstable RLO (Tauris et al., 2011) in which the expanded Roche Lobe of the donor completely engulfs the NS in its hydrogen rich environment, leading to a common envelope (CE) phase (Paczyński, 1976) and subsequent inspiraling of the donor. This most likely causes a merger, likely resulting in either an isolated NS or a black hole. LMXBs in some cases can also undergo a late Case B RLO, leading to a very wide ($P_\mathrm{b} > 1000$ days) NS-CO WD system with slow ($P \sim 1$ second) pulsar and a WD companion mass of $0.47 - 0.67\,M_\odot$ (Tauris et al., 2012).

- *Case C* RLO occurs if a relatively massive ($\sim 3\,M_\odot$) donor is at or beyond the core helium burning stage of its life and is clmbing the giant branch for a second time. This is rarely observed in comparison to Case A or B RLO as it requires the pre-mass transfer progenitor system to have an extremely wide orbit (P_b $100 - 1000$ days). Case C RLO describes the situation where the initial Case A or B RLO triggers a 'chain reaction' of accretion to the NS followed by orbital shrinking due to mass-loss.

[13] Some isolated pulsars are theorised to have once been in binary systems (see e.g. McLaughlin et al., 2005)

This leads to further RLO and, eventually, an unstable mass transfer followed by a CE. Within the CE, the inspiraling NS causes large dynamical friction and, in a wide binary ($P_\mathrm{b} \gtrsim 1$ yr), often leads to severe loss of angular momentum and ejection of the envelope. Many potential binary systems merge during the unstable CE phase (Glebbeek et al., 2013) as in Case B RLO; however, for Case C RLO, the CE is less tightly bound than in Case B so there is a greater chance that the MS will survive the inspiral phase. Those that survive as binaries go on to form a NS–He (neutron star – helium star) system, where the He-star is the naked He–burning core from the former MS companion. The final result of NS-He binaries being possible DNS progenitors is outlined below.

Low and intermediate XBs that result in NS-WD systems do so because the MS companion does not have enough mass to end its life in a core-collapse and supernova and instead moves up the red giant branch eventually becoming a white dwarf. On the other hand, HMXBs that have undergone Case B or C RLO after the first supernova may contain MS companions that are massive enough to result in a supernova. If a sufficiently massive NS–He binary is in a close orbit, a second RLO may occur. In the majority of cases this is achieved via Case BB RLO (Tauris et al., 2015; Tauris et al., 2017) if the donor mass has less than approximately $3\,M_\odot$ following Case B RLO. This will be a lengthy and stable accretion process through the L_1 Lagrangian point. Case BB RLO causes the NS to "spin–up" due to its now increased angular momentum, sometimes to rotation periods of milliseconds to tens of milliseconds (Alpar et al., 1982). This also sees a decrease in magnetic field strength by up to 5 orders of magnitude, possibly due to the accretion process (Bisnovatyi-Kogan and Komberg, 1974), which would explain the low magnetic field strengths observed for MSPs. The larger NS rotation periods observed for DNS systems in comparison to NS-WDs imply that recycling is not as efficient for HMXBs as it is for LMXBs. Globular clusters contain a dense population of massive stars and thus many pulsars observed in them are in binary systems with extremely low spin periods (see Jacoby et al., 2006, for a discussion about the DNS B2127+11C). The evolution for binaries in globular clusters, where Case C RLO is expected more often leads to long-period blue stragglers. This is rather different from Galactic binary evolution and will not be discussed further.

The Case BB RLO may have the effect of tidally stripping the companion He star if the orbital separation is small enough. If the XB results in a late, deeply convective, Case B RLO, and the post-CE He core is less than $\sim 1.375\,M_\odot$, the core will become an O-Ne WD (Postnov and Yungelson, 2014) and the resulting system will be a NS-O-Ne WD binary. However if the progenitor IMXB/HMXB results in an early, radiative, Case B RLO and minimal mass loss occurs during the relatively stable Case BB RLO, eventually either free electrons are rapdily captured by ^{24}Mg and ^{20}Ne in the former He-star core (ECSN; Miyaji et al., 1980) or, for more massive companions (which are more likely to form a NS instead of a WD), the ultra-stripped iron core undergoes rapid core-collapse (FeCCSN; Tauris et al., 2013), eventually leading the star to go supernova. This results in another *young*

NS[14]. The new system comprises two NSs and is thus called a *double neutron star* (DNS) system. This does not explain highly eccentric systems with high transverse velocity, such as PSRs B1913+16 and B1534+12, in which the companions are thought to have resulted from a Type II SN.

Because recycled NSs typically have much weaker magnetic fields than young NSs, based on a purely dipole emission model, they have much longer observable lifetimes which explains why all but two pulsars in Galactic DNS population are known to be recycled. The two exceptions are PSR J1906+0746 (P = 144.1 ms; van Leeuwen et al., 2015), in which the companion is likely the recycled NS, not seen due to beam geometry[15]; and the companion (B) star of the "double pulsar" (PSR J0737−3039A/B, $P_{\rm A}$ = 22.70 ms, $P_{\rm B}$ = 2.773 s; Burgay et al., 2003), so called because both of its constituents were discovered as pulsars. Observations of DNS systems for which the pulsar is the recycled NS have shown a strong correlation between the measured space velocity and the mass of the second-formed (companion) NS. It is believed that DNSs with lower space velocities ($< 100\,{\rm km\,s^{-1}}$) developed from symmetric second supernovae which imparted a small natal kick onto the system following Case BB RLO (Tauris et al., 2017), usually resulting in relatively low eccentricities. DNSs with higher space velocity and eccentricity are likely to have formed in an asymmetric second supernova which imparted a large kick, as is observed in the estimated space velocities of PSRs B1913+16 ($v^{\rm LSR} = 150\,{\rm km\,s^{-1}}$) and B1534+12 ($v^{\rm LSR} = 120\,{\rm km\,s^{-1}}$; Haniewicz et al., 2021). Another observed correlation is between the spin period of the recycled NS and the orbital eccentricity (McLaughlin et al., 2005). The correlations concerning eccentricity are presumably both a consequence of the slow Case BB RLO mass transfer phase[16], in which the stable mass transfer would result in lower mass loss during the second supernova, leading to a lower eccentricity if the second supernova is symmetric. The observed correlation between the eccentricity with spin and companion mass could also be due to a selection effect favoring low-P/low-e systems because more eccentric close binaries will merge quicker (Chaurasia and Bailes, 2005) and therefore no longer be observable. Globular cluster pulsars, pulsars in triple systems (e.g. PSR J0337+1715; Archibald et al., 2018), as well as pulsar–planet systems (Sigurdsson, 2003) do not fit these evolutionary scenarios well due to the added dynamical complexities arising in these environments.

With so many types of stellar remnant systems all having undergone slightly different evolution scenarios, it is vital that observations of pulsars in binaries be made to high precision in order to best constrain the evolutionary parameters mentioned.

[14]Another proposed supernova theory for DNSs concerning a double core evolution is given by Dewi et al. (2006)

[15]It is not ruled out that this system might be a NS-WD.

[16]The longer the transfer phase, the more spun–up the recycled NS will be and the less massive the He star will be.

1.4 This thesis

This thesis is organised into three main chapters which highlight the range and depth of pulsar timing studies.

In Chapter 2, the data collection and down stream signal processing methods are discussed. New methods of RFI excision via neural network image recognition are compared with traditional methods such as sigma–clipping, which are often biased due to inevitable TOA component mixing in the resulting residuals. Methods concerning signal calibration are also explored.

Chapter 3 documents the most up-to-date timing analysis of a pulsar in a relativistic DNS system: **PSR J1829+2456**, where significant measurements have been made for both the component masses and the system proper motion. The chapter concludes with a discussion of the evolution of the system based on the new timing solution.

In Chapter 4, four pulsars recently discovered by the PALFA pulsar survey are investigated: **PSRs J1851+0010**, **J1853+0008**, **J1936+1805** and **J1936+2142**, three of which are believed to be in binary systems. This chapter presents the best current timing solution and physical interpretation for each pulsar.

Chapter 2

Observation setup and data processing

Observations made for the purposes of this thesis were all taken with the spherical 305-m primary dish at the Arecibo Observatory (AO) in Puerto Rico, and using the Puerto Rico Ultimate Pulsar Processing Instrument (PUPPI) backend. Although the telescope is now due to be demolished, at the time of writing the basic observational setup and metadata for the telescope's receivers[1] and the backend[2] can be found on the AO website. A brief description of the specific setup for AO observations is given here:

1. Each of the ten receivers at AO have two dipole feeds, set up at orthogonal (for linear) or opposite-handed (for circular) bases. Data for this thesis was taken with the 430 MHz and L–wide receivers, with the dipole feeds set to be dual linear. In reality, pulsar emission is elliptically polarised but by characterising the beam in terms of the Stokes parameters, one can determine the total beam intensity which determines the total power output, so it does not matter which orthogonal basis is used. Polarisation at the telescope will be expanded upon in Section 2.3.

2. Raw radio signal from the sky at the dipole feeds (call this $\nu_{\rm src}$) is down–converted to an intermediate frequency (IF – $\nu_{\rm IF}$) signal to reduce internal signal loss due to pipeline elements (cables, etc.). This signal loss is roughly proportional to the square of the signal frequency, so that a lower frequency suffers less overall loss. The telescope at the AO uses a superheterodyne system (Armstrong, 1921) to mix $\nu_{\rm src}$ with a monochromatic local oscillator (LO) of frequency $\nu_{\rm LO}$ set to the Hydrogen Maser Frequency Standard of 1,420,405,751.77 Hz. This creates a high–mix frequency and a low–mix frequency, shown in Equation 2.1 as the sum and difference of the frequencies respectively.

$$\begin{aligned} \nu_{\rm IF} = \nu_{\rm src}\nu_{\rm LO} &\approx \sin\left(\omega_{\rm src}t\right)\sin\left(\omega_{\rm LO}t\right) \\ &\sim \cos\left(\left[\omega_{\rm src}-\omega_{\rm LO}\right]t\right) + \cos\left(\left[\omega_{\rm src}+\omega_{\rm LO}\right]t\right) \end{aligned} \tag{2.1}$$

[1] http://www.naic.edu/~astro/RXstatus/rcvrtabz.shtml
[2] https://www.cv.nrao.edu/~pdemores/puppi/

3. Once mixed, the IF signal is put through a low–pass[3] filter to obtain the signal relative to the LO's frequency. Stages 2 and 3 are collectively known as the IF/LO chain[4].

4. The new frequency at $\nu_{\rm IF}$ was then digitised and passed to the PUPPI backend where further processing, such as online dedispersion and profile folding, took place.

2.1 Initial data processing

The datasets were downloaded to the local UEA file server from Arecibo's 'puppimaster' server and with some having been obtained from the servers at McGill University. For this work, a combination of both incoherent search-mode data and coherently dedispersed fold-mode data were used to obtain timing solutions for five pulsars. The search-mode data were first divided into subbands of 256 or 512 frequency channels and initial times-of-arrival (TOAs) were obtained through the `get_toas` routine in the PRESTO pipeline. This involved two main steps: a rough method for radio frequency interference (RFI) excision (`rfifind`), in which the time series was divided into segments of 0.5 s and each segment analysed for RFI; followed by a folding routine (`prepfold`) set to the best-fit folding period for each session of observation (see Ransom, 2001, Section 5.1) at a profile resolution of 128 or 256 bins. These incoherent TOAs were then *phase connected* iteratively using TEMPO[5] by removing artificial constant phase jumps (manually input phase corrections designed to artifiially align the pulse phases) between sets of different observing epochs and fitting physically motivated parameters as necessary until all the TOAs could be described by a spin model that was free from inter-epoch phase offsets. This is called a *phase-coherent solution* and takes into account every rotation of the NS, assuming that parameter measurements do not contribute to integer numbers of rotation period offsets. Once phase-connected, this initial phase-coherent solution is called the *initial solution*. To obtain the solutions as reported in this work, the search and fold-mode data were refolded using the initial solution to achieve more accurate TOAs, and therefore a more precise solution. Re-folding the data with a correct ephemeris can also bring out more nuanced features in the pulse profile and is often done iteratively over older data as new data becomes available after the initial discovery.

During refolding, the raw search-mode data was directly folded with the initial solution using the `fold_psrfits` routine in the `psrfits_utils` package[6], with a minimum of 256 bins per profile. These now-folded data, together with the other fold-mode data, were analysed using the PSRVoid pipeline, which is now described below and further detailed in Appendix C, in order to remove further RFI contamination.

[3] Radio astronomers use low/high–pass to refer to frequency, whereas optical astronomers refer to wavelengths. Here, we obviously use the radio nomenclature.

[4] A full set of schematics for the entire IF/LO chain used at the Arecibo Observatory is available at http://www.naic.edu/~astro/techinfo/iflo/

[5] http://tempo.sourceforge.net

[6] https://github.com/demorest/psrfits_utils

2.2 RFI excision

The ability to make accurate and reliable TOA measurements is dependent on the quality of profiles that are averaged to form the integrated profile. During an observation, data is sometimes overshadowed by noise and unwanted signal which can be broadly categorised into three main areas of interest:

- Instrumental (or system) noise, which is due to system components such as amplifier setup and receiver component temperature (e.g. see above).
- Astrophysical noise, which describes any noise source that is astrophysical in origin. This category includes red timing noise due to unmodelled physics within the data (Hobbs et al., 2006b), the cosmic microwave background noise, and other porminent local backgrounds.
- Interference, which is due to unwanted signals (of usually unknown origin) degrading the quality of the true signal. In astrophysics, interference is most often in the form of electromagnetic interference, and for radio astronomy: RFI.

RFI is a common source of interference in pulsar data and can be frequency specific (irreconcilable at a certain frequency over the entire observation), or epoch specific (happening over a short burst of time). Narrowband RFI is interference that takes place over (theoretically) one frequency channel and is often due to small communication devices such as mobile phones. Broadband RFI exists over a much larger range of frequency channels and, whilst often at a lower amplitude, is more difficult to fully disentangle from the signal than in the case of narrowband RFI. After averaging, RFI in unclean datasets can sometimes reduce an otherwise good signal to a null profile (a profile that is 'flat' and often close to 0 over the phase-space), and other times narrowband RFI far exceeds the pulse flux density and the only visible signal is that from the interference source. This can be explained by receiver *saturation* in which the limiting output voltage of the analog receiver is exceeded by the interference source power input. Profiles such as this are said to saturate the radio source, making it undetectable until the interference is removed (or clipped). An added caveat with saturated profiles are that, due to harmonics created by clipping the RFI regions, even narrow-band RFI can affect the total input frequency band. Including especially these latter types of profiles when refolding may shift the integated profile maximum by some non-negligible fraction, producing TOAs that no longer effectively model the emission or rotation. The effective output of the routines described in this section is an ASCII file containing a list of pairs of sub-integrations and frequencies, corresponding to individual polarisation–averaged profiles, which meet a specific criterion for flagging the profile for removal due to RFI.

Removing RFI from folded profiles can sometimes be 'too late' in the sense that the RFI is persistent throughout the entire band (due to harmonics, etc.). In this case, small chunks of time on the order of a few hundred ms may need to be removed. In the most extreme cases, entire bands may need to be abandoned for a particular observing day, although this is rare.

In this thesis, three models were exploited for identifying RFI in folded pulsar data and developed in code: Gaussian σ–clipping, image recognition via machine learning, and overtone matching. The models were then compared against each other by using the timing solutions resulting from the TOAs produced with the corresponding integrated profiles (see Section 2.2.4).

2.2.1 Gaussian σ–clipping

The most simplistic model used to identify RFI assumes a box–like structure for the on and off-pulse regions as determined by a template profile and further assumes that mean flux density RMS in the off-pulse region follows roughly a Gaussian distribution (or a set of Gaussian distributions). These assumptions are valid if care is taken to remove known sources of RFI; if there is an extreme bias across multiple channels, the off-pulse signal will often not be well modelled by a Gaussian. In these clear-cut cases, the known sources of RFI can be manually removed before further processing. σ–clipping works by removing profiles that lie outside some multiple of the standard deviation for some given statistic (here that is the off-pulse RMS). There are two fundamentally different statistical criteria that the PSRVoid package uses for σ–clipping: Chauvenet's criterion (Chauvenet, 1863) and double median absolute deviation (DMAD), which are described below.

Chauvenet's criterion assumes a single Gaussian, with mean μ_{rms}, and is the least computationally expensive of these two methods. The mean RMS for N profiles, from the frame of reference of a whole profile, x_i, is given by:

$$\mu_{\text{rms}} = \frac{1}{N} \sum_{i}^{N} w_i x_{i,\text{rms-opw}} \tag{2.2}$$

where w_i is the frequency density of the RMS value for the i^{th} profile and the OPW subscript stands for the off-pulse window: the aforementioned 'box–like' region as determined by profile bins where the template profile is identically 0. Chauvenet's criterion suggests that any profiles whose specific off-pulse RMS is not in-between $\pm a\sigma$, for some user-defined tolerance a, is either a null profile at the lower extreme or a high-flux density affected channel due to narrowband RFI at the higher extreme. Thus, any profile that is outside this range is given a weight of zero for further analysis. To the keep the same resolution of significance throughout the timing process, RMS values outside the range for $a = 2$, which corresponds to a 95% Gaussian confidence, were rejected. Considering the number of profiles per observing day (~ 60000), this was decided as an appropirate limit.

The second approach, DMAD, assumes that prominent sources of interference are often independent of each other and therefore may exhibit local Gaussian flux density distributions; however these local interference distributions will generally not comprise a global Gaussian flux density distribution. This is generally true and is not taken into account in Chauvenet's criterion. Taking a single value of μ_{rms} for a multi-peaked interference distribution using Equation 2.2 may, in the most extreme cases, omit every profile due to the $a\sigma$ cut-off boundaries being unrepresentative of the data. DMAD instead calculates the median of each local distribution (or more practically, the n most prevalent local dis-

tributions) and compares pairs of medians via calculation of the modified Z-score (Iglewicz and Hoaglin, 1993), which is a standardised score that measures outlier strength from the median in units of standard deviation. For an interference distribution with n Gaussian signals, the modified Z-score for the i^{th} profile, Z_i, is given by:

$$Z_i = 0.6745(x_i - \bar{x}) \left(\sum_{j}^{n} |\overline{x_i - \bar{x}_j}| \right)^{-1} \tag{2.3}$$

where a bar over an element represents its median and $x_i \equiv x_{i,\text{rms-opw}}$ from above. The modified Z-score is compared against a user-defined threshold, with modified Z-score values less than the threshold signifying a strong outlier signal, among which RFI is expected to lie. In effect, this determines how the local distribution compares with respect to the overall RMS population (Leys et al., 2013). If the local distribution is centred around a highly noisy source, then the denominator of Equation 2.3 will be large, making it less likely that the resulting modified Z-score will exceed the imposed threshold, and thus more likely to be classified as an outlier interference source (such as RFI). Iglewicz and Hoaglin (1993) suggest a threshold value of 3.5 for 95% confidence (for a boundless number of samples) so this value has been adopted for this work. This method also somewhat removes the need to manually omit known sources of RFI beforehand, as was the case using a single Gaussian approach (assuming that the known source of RFI is itself roughly Gaussian-distributed). Although the DMAD Python implementation is slower than using Chauvenet's criterion by about 25% for the case of a two-component off-pulse RMS distribution (the theoretical best case scenario) and significantly slower for more sensitive DMAD tests, the time saved in reducing manual intervention makes this a faster method for any multi-Gaussian interference distributions.

2.2.2 Image recognition excision via deep learning

Gaussian σ–clipping can rapidly result in broad statements about large data volumes, however RFI–riddled 'inlier' profiles may still persist if their off-pulse RMS coinicidentally lies within the acceptable threshold, despite not contributing to the true pulsar signal in a meaningful way. Dedispersed L–band data for PSR J1829+2456 taken for this thesis (see Chapter 3) contained several profiles that displayed this property, mainly residing on the outskirt edges of particularly noisy sub-bands. One solution to this might be to increase the cut-off threshold and concede to a greater measurement uncertainity; however, to preserve the 95% significance tests used in this work, a different approach was necessary.

Machine learning (ML) for pattern recognition has already been used for image and speech recognition on platforms such as Facebook (Pinheiro et al., 2015; Pinheiro et al., 2016; Baevski et al., 2020), and is a growing topic in the astrophysical community. This includes pulsar astronomy, for wide reaching studies, such as profile evolution and emission in particularly complex situations (e.g. Devine et al., 2016; Kwofie, 2018), as well for pulsar candidate searches (Zhu et al., 2014). Given the success of ML in these instances, it has been explored here as a potential improvement on Gaussian σ–clipping in order to detect and excise RFI.

Machine learning (Samuel, 1959) represents the defiance of 'conventional programming', in which the behavior of the algorithm is a function of the input variables, as opposed to a set of static protocols programmed by a user, in order to predict patterns in a way that a human programmer could not, and therefore could not programme for. Classical ML has often been a process of a feature-matching. In this type of ML, a predetermined set of macro-features from an image, spanning more than one pixel, is preprogrammed as a validator to compare to the input pixels and decide if the input matches any known macro-feature (e.g. Taylor et al., 2009). This is generally the least computationally expensive ML approach so is preferred for well-understood images and patterns; however the often ill-understood sources of interference in PSRFITS data require an approach that is able to determine relevant features without already having access to a look-up table. In contrast, modern ML must be trained with 'live' examples. This involves creating a training set, in which the outcomes from each example in the set are known, and comparing that with the responses given by the artificial intelligence (AI) algorithm for the same set. If the AI's setup caused it to categorise many examples incorrectly the AI changes its behavior, based on metrics outlined below, until its responses better match those of the training set's outcomes. Since the training set can be any sequence of data points and the rest of the algorithm evolves through training, modern ML algorithms can adapt to many situations and are therefore more economical than classical systems overall. This makes them better suited than feature-matching ML algorithms to the problem of classifying pulse profiles, which requires a great deal of flexibility.

There are many radically different types of infrastructure for modern ML algorithms which have been explored in this work. Two main competing types of ML algorithms are the recurrent neural network (RNN), an infinite impulse response network useful for analysing analog signals where the continuous nature of the data does not require a zero-time, and the convolutional neural network (CNN; neural network – NN) (LeCun et al., 1989), a finite impulse response network designed to mimic the neuron pathways in the human brain for discrete (digital) systems. As such, CNNs were chosen for analysing the digitised pulsar data. CNNs are also sometimes known as *shift invariant* NNs owing to their shared-weight architecture, as will be discussed in this chapter. An excellent overview of CNNs is given by Goodfellow et al. (2016, Chapter 9). CNNs are often easier to design than RNNs when particular responses are known (such as known patterns or desired responses) but RNNs are generally more efficient to implement owing to a smaller list of initial assumptions.

Abstractly, neurons in a CNN that fire action potentials more often for certain inputs end up end up being 'reinforced' when presented with that those inputs multiple times. Conversely, neurons that fire less over the same experiences are not nurtured and become undermined. Thus, when a comparable situation arises again, the neurons fire those potentials more efficiently due to stablised and strengthed neuron-to-consequence mapping[7]. This can be represented by the input vector and a series of one-dimensional vectors, known as *layers*, comprising an output layer and an intermediate set of *hidden* layers which are mathematically coupled to each of the input and output nodes, as shown in Figure 2.1.

[7]This is how the human brain learns, but it is currently impossible to replicate the complixity of a brain.

The connections between one node in a layer and a node in an adjacent layer are called *folds*[8].

During the feed-forward process, there will eventually and necessarily be a reduction in the number of nodes. If there is not, and the number of categories equals the number of input nodes, the CNN will not be able to associate regions as effectively. This node reduction can be seen as a filter for the input, simplfying it down to more rudimentary patterns, until one value is output as the test result. The output volume can be controlled via three hyperparameters: the *depth*, the *stride* and the *zero-padding* – outlined below:

- The depth, n, is the connective scope of the NN architecture. As a CNN learns from analysing feed-forward results, recurrent features in the input space, when seen by the AI, begin to 'fire' recurrent patterns through the layers. The longer the chain of layers, the more neurons there are in your system holding information about the input space and the more complex that optimal feature search can become. In most cases, depth can be thought of as the number of layers, with a *depth slice*, or node, denoting a value in a given layer.

- The stride, s, determines the vector length change from layer to layer. In the $n = 3$ case, the stride of the only hidden layer in a CNN is the same as the search window used in Gaussian σ-clipping. This is because all points within the search window are averaged in Gaussian σ-clipping, with is much the same process here. For a deep CNN, the stride determines the speed at which the network learns per layer, although this is necessarily a mean value.

- The zero-padding parameter, $q \equiv s/2$ is used to compensate for clustering effects due to averaging at image edges with constant stride. For a layer-layer connection of stride s, the first and last q depth slices in the first layer will each have less values over which to average than the other values in that layer. The effect is a cumulative mean in those particular depth slices meaning that the first and last few nodes in the resulting layer will not be mutually exclusive which leads to biased weightings between those layers. To overcome this, q null values are applied to either side of the image. This is not suitable for every layer if the exact spatial size of the input image is important to preserve.

Generally, the feed-forward output, $\mathbf{Y}$, for an n–layer NN can be written recursively in terms of the input nodes, $\mathbf{X}$, the weight matrices, $\mathbf{W}^{(i)}$ (for layer i) and bias nodes, $\mathbf{b}^{(i)}$, which are preprogrammed constant offsets to apply to nodes in a given layer, as described below:

$$\mathbf{Y} = f_n\left(\mathbf{W}^{(n)} f_{n-1}(\mathbf{Z}^{(n-1)}) + \mathbf{b}^{(n)}\right) \tag{2.4}$$

where f_i is some non-linear function designed to model non-linear behavior in the output, and

[8]Not to be confused with the pulsar astronomy definition of folding!

$$\mathbf{Z}^{(i)} = \mathbf{W}^{(i)} f_{i-1}(\mathbf{Z}^{(i-1)}) + \mathbf{b}^{(i)} \tag{2.5}$$

with

$$\mathbf{Z}^{(1)} = \mathbf{W}^{(1)} \mathbf{X}^{\mathrm{T}} + \mathbf{b}^{(1)} \tag{2.6}$$

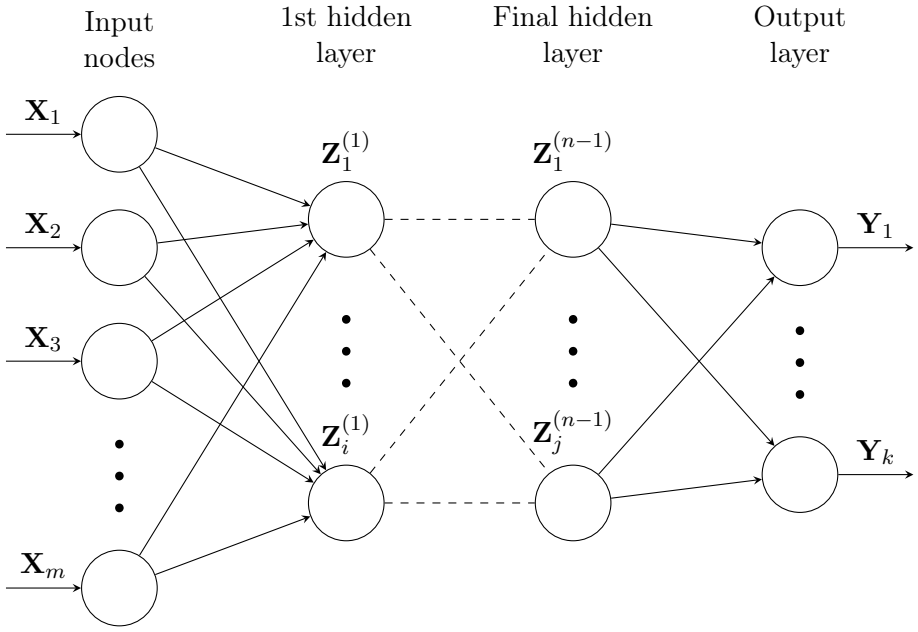


Figure 2.1: A general overview for a deep-learning, n–layer neural network. Each layer has a shape independent of the other layers (i.e. in general, $m \neq i \neq j \neq k$, etc.). Layers are all linked together by (generally non-square) weight and bias matrices. The dashed lines represent potentially many hidden layers connected linearly. The input nodes do not count as a layer. For example, the input layer is the individual data points in a pulse profile and each subsequent layer is, in effect, a binning-down of these data points, not too dissimilar from classic smoothing procedures.

The term 'deep learning' is most often attributed to NNs where the number of hidden layers > 1, as it has been shown that increasing the number of hidden layers increases accuracy, so is frequently conflated with a higher degree of learning. Although it is possible to set specific initial weights, the point of using CNNs is to be able to start from essentially a random uniform distribution of weights and create 'order from the chaos' by way of 'back-propagation'. The training is achieved via two training sets: the normal training set, $(\mathbf{X}_{\mathrm{tr}}, \mathbf{Y}_{\mathrm{tr}})$, and the validation set, $(\mathbf{X}_{\mathrm{val}}, \mathbf{Y}_{\mathrm{val}})$, which contains profiles that are representative of the normal training set population but independent of that set and generally contains only 20% of the content, a split determined by minimizing the validation and training set error rates (Guyon, 1997). In general, the number of validation set profiles should be roughly inversely proportional to the square-root of the number of free parameters. These training sets contain individual pulse profiles as inputs with a key to designate them as either containing RFI (0) or clean (1). Training, then, is a process of reading the training set inputs, applying the initially 'random' weights, comparing the output result with the known training output and back-propagating down the neurons in order to update

the weights if necessary, to better match the expected output result. A final metric used for effective NN design is the Vapnik-Chervonenkis dimension (Vapnik and Chervonenkis, 1971), which is the overall measure of NN complexity and will be discussed later in the chapter.

Loss

A further set of values representing the current rate of success, the *loss*, is also calculated. The loss is an important property which determines how well the NN learns over many iterations with respect to a given *loss function*[9]. For binary classification, the most widely used loss function is the binary categorical cross-entropy (BCCE) function, which works by comparing the output node values with each other and then assigning the input nodes as belonging to one or the other category based on a given threshold for one. This means that it strongly penalises confident yet wrong predictions. PSRVoid (Appendix C) implements BCCE, for loss $\mathcal{L}$, and a given training set (signified by the subscript 'tr') as follows:

$$\mathcal{L} = -\frac{1}{N_{\rm tr}} \left(\mathbf{Y}_{\rm tr} \cdot [\log \mathbf{Y}]^{\rm T} + (1 - \mathbf{Y}) \cdot [1 - \log \mathbf{Y}_{\rm tr}]^{\rm T} \right) \tag{2.7}$$

where $N_{\rm tr}$ is the number of samples in the training set. The '1's are identity matrices with the same shape as $\mathbf{Y}$. This loss function is identical to the one calculated for the validation set with subscript 'tr' replaced with 'val'.

Amending the weights during back-propagation is conducted by a process of 'steepest gradient descent' (SGD[10]) analysis (Curry, 1944) with the goal of minimizing the loss function on the typically multi–dimensional loss surface. The updated position, $\mathbf{p}_{i+1}$, of the point in loss-space at step $i+1$, can then be expressed in terms of an arbitrary *learning rate*, l, as:

$$\mathbf{p}_{i+1} = \mathbf{p}_i + l \nabla g \tag{2.8}$$

for arbitrary function g whose gradient is defined as being negative. The learning rate is a positive scalar that codifies the loss-space step-size of the gradient search from one point to the next. A learning rate that is too large may pass over narrow but deep troughs, whereas a NN with a learning rate that is too small may get stuck in a broad but shallow trough and therefore stop improving. In this work, l is kept constant and small ($\sim 0.001 - 0.01$) to achieve a good resolution-to-time ratio; however, gradient line-search methods in which the learning rate is also a function of the current gradient (e.g. Nocedal and Wright, 2006) can be used to determine more optimal learning rate values at each iteration. This is was not implemented for testing the models for RFI excision, as line-search implementation on a subset of the data made little difference on the quality of the NN outputs for a given training set over a constant l method[11].

[9] Also called the *objective function* (cf. Goodfellow et al., 2016, p. 80)

[10] The NN community use SGD to mean 'stochastic gradient descent' (cf. Goodfellow et al., 2016) which involves splitting the data up into sub-sections to optimize large NNs. This work does not require such optimisations and they are not used, so the term SGD stands for 'steepest gradient descent'.

[11] The line search takes a lot longer to run however.

Table 2.1: Properties of various activation functions used for testing, including their derivatives, ranges and descriptions.

Name	$f(x)$	$f'(x)$	Range	Description
Sigmoid $\sigma(x)$	$(1+e^{-x})^{-1}$	$\sigma(x)(1-\sigma(x))$	$(0,1)$	Useful for binary classification. Gradient is gentle but can flatten altogether. Computationally expensive.
Hyperbolic tangent $\tanh(x)$	$2(1+e^{-2x})^{-1}-1$	$1-\tanh^2(x)$	$(-1,1)$	Zero-centred, meaning better classification for strong negative values. Otherwise, it is a scaled Sigmoid.
Rectified linear unit*, ReLU(x)	$\max(0,x)$	$\begin{cases} 0 & x<0 \\ \text{undef} & x=0 \\ 1 & x>0 \end{cases}$	$[0,\infty)$	Computationally efficient. Back-propagation fails if input approaches 0 because of the undefined derivative.
Softmax† $\mathrm{s}(\mathbf{x})$	$e^{\mathbf{x}}\left(\sum_{j=1}^{N} e^{x_j}\right)^{-1}$	$\frac{\partial \mathrm{s}_i}{\partial q_j} = \mathrm{s}_i(\delta_{ij}-\mathrm{s}_j)$	$(0,1)$	Normalises input vector, $\mathbf{x}$. Can be used for classification of more than two categories.
Swish $\varsigma(x)$	$x\sigma(bx)$	$\varsigma(x)+\sigma(bx)(1-\varsigma(x))$	$[\approx -0.278/b, \infty)$	Smooth throughout entire domain. b is a constant scalar.

*A slight modification to ReLU is the Leaky ReLU function (LReLU$(x) = \max(ax, x)$), in which the scalar, a, can be fine tuned to model negative value behavior in the input, however results from LReLU–driven nodes vary wildy depending on the nature of the negative values.
†It makes no sense to talk about the Softmax function with an isolated node, x, within a fold as can be done with the other activation functions mentioned here. Therefore, the vector $\mathbf{x} = \{x \mid x \in \mathbf{Z}^{(j)}\}$ is used as the argument instead.

Activation functions

The set of functions in Equation 2.4, $\{f_i\}$, are known as activation functions and can be categorised into three main groups: binary step functions, where the input is compared to a threshold and the neuron is activated if the criteria is met; linear functions, where the output is a linear combination of the input; and non-linear functions. The first two of these groups are unsuitable for deep learning; binary step activation functions cannot support mutliple categories of data, and linear activation functions are both always collapsible to a 2–layer network and cannot support back-propagation due to the constant nature of the derivative (i.e. there is no 'steepest gradient'). The non-linear activation functions are the only choice for back-propagated deep learning algorithms, as these can model the generally non-linear behavior of neurons with a linear system of equations (e.g. Equation 2.5).

The choice of non-linear activation functions is dependent on the particular design of the NN, with some functions being preferable at certain layers. For the purposes of this work, the choices for the functions were decided based mainly on qualitative grounds from the successes of many previous studies, however recent research has yielded promising results for determining the optimal type of activation function via a ML approach (Manessi and Rozza, 2018; Nader and Azar, 2020). There are a vast collection of optimised activation functions for specific types of NN, five of which are explored for this thesis (Table 2.1[12]), often (and here) quoted as functions of one node, x, unless stated otherwise.

To determine the most effective activation architecture for the RFI excision NN, six unique 4–layer CNNs were constructed with node structure $(|\mathbf{X}|, 32, 16, 32, 2)$ where each

[12]For a full derivation of the derivatives, see Appendix A.3

number is the number of nodes in that layer. Weight matrix elements were all initialised with NumPy's (van der Walt et al., 2011) random seed generator with a seed of '1' (although this choice is arbitrary), and bias elements all set to 0. The acceptance tolerance, the value that the output must be equal to or greater than to be labelled as 'clean', was set at 0.7 although lower tolerances were also explored. The first activation function in the feed-forward was chosen to be either the ReLU or the Swish function, the second function was kept fixed on the ReLU, and the output layer was determined by either the sigmoid, the hyperbolic tangent or the softmax function. The positional choices for these activation functions were based on each function's specific ability to handle the incoming data: the initial hypothesis was that the Swish + sigmoid approach would produce the 'smartest' and least time-demanding NNs. This is due to Swish's ability to negotiate negative values from the input layer instead of simply nullifying them, whilst otherwise behaving mostly like the ReLU function, and sigmoid's effectiveness at binary classification. The softmax normalisation property helps to handle vanishing gradients present in the output layer, but may have a difficult time with binary classifications due to limited input nodes when back-propgating. The softmax function can be used to more easily compare data on different scales[13].

The training data for this experiment was created from pre-averaged PUPPI observations of PSR J1829+2456 at 430 MHz taken on MJDs 58402, 58404, 58406 and 58408. This made for a total of 375 profiles in the training set and 75 profiles in the validation set. Within each set, 20 null profiles with their output flag set to 0 were included in order for the networks to more easily learn the structure of null profiles. During training, the loss was monitored over $100,000$ iterations, or until convergence to four decimal places, for each CNN at six different learning rates. The final loss values are recorded in Table 2.2 and the full loss profiles are given in Appendix B.2. For effective and realistic learning, the loss profile should be a smooth decrease with increasing iterations, however some discontinuities remain and appear more frequently as a function of increasing l. This is a result of the learning rate being too high to stick to a particular potential well in the $\mathcal{L}$-landscape – although, the ReLU/tanh CNN structure produced an excellent learner at $l = 0.0003$ despite the loss profile's deviation from smoothness.

A *confusion matrix* (Stehman, 1997), C_{ij} (for $i, j \in \{0, 1, ..., n\}$), was then calculated using the Scikit-learn Python package (Pedregosa et al., 2011) – C_{ij} visually represents the number of correct and incorrect predictions made by the weight and bias matrices obtained from the final iteration of training. The confusion matrix has four defined performance metrics (Kwofie, 2018) for which higher values all-round suggest a more effective NN:

- The *accuracy*, the fraction of all correct predictions,
- The *sensitivity* (also called the 'true positive rate'), the fraction of clean profiles that were correctly predicted,
- The *specificity* (also called the 'true negative rate'), the fraction of all noisy profiles that were correctly identified,

[13] A common occurence is the conversion from counts to Jy (see Section 2.3 for a different approach).

- The *precision*, the probability of a correct prediction given that the prediction was a 1 (clean profile).

These metrics, for a binary output system, are outlined as functions of C_{ij} in Equations 2.9a – 2.9d respectively. The four metrics were calculated for each activation setup at the specific learning rate which produced the minimum loss for that activation setup (Table 2.3 – see Table 2.2 for the learning rates).

Table 2.2: The loss for each unique 4–layer CNN after 100, 000 iterations, or less depending on convergence, of training at six different constant learning rates. The second activation function (always ReLU) is implied and therefore omitted from the setup, so the setup is quoted for the first and third functions in that order. The ReLU function is here denoted $\mathrm{R}(x)$.

	Loss, $\mathcal{L}$	Learning rate, l					
		0.0001	0.0002	0.0003	0.0004	0.0005	0.001
Activation setup	$\mathrm{R}(x), \sigma(x)$	0.1437	0.1309	0.1240	0.1166	0.1088	0.0954
	$\varsigma(x), \sigma(x)$	0.1411	0.1296	0.1223	0.1143	0.1065	0.0968
	$\mathrm{R}(x), \tanh(x)$	0.1318	0.1334	0.1333	0.1320	0.1484	0.1328
	$\varsigma(x), \tanh(x)$	0.1408	0.1203	0.1211	0.1374	0.1465	0.132
	$\mathrm{R}(x), \mathrm{s}(\mathbf{x})$	1.9670	1.9689	1.9774	1.9774	1.9774	1.9647
	$\varsigma(x), \mathrm{s}(\mathbf{x})$	1.9542	1.9767	1.9493	1.9492	1.9493	1.9492

Table 2.3: The performance metrics (in %) for the 4–layer CNN as calculated from Equations 2.9a – 2.9d for six types of activation architectures.

Architecture	$l\,(\times 10^{-4})$	Accuracy	Sensitivity	Specificity	Precision
$\mathrm{R}(x), \sigma(x)$	1	97.9	93.6	100	100
$\varsigma(x), \sigma(x)$	1	97.9	93.6	100	100
$\mathrm{R}(x), \tanh(x)$	3	95.7	87.2	100	100
$\varsigma(x), \tanh(x)$	3	96.5	89.6	100	100
$\mathrm{R}(x), \mathrm{s}(\mathbf{x})$	1	66.7	0	100	0
$\varsigma(x), \mathrm{s}(\mathbf{x})$	4	77.6	32.8	100	100

$$\text{Accuracy} = \left(\sum_{i=j} C_{ij}\right)\left(\sum_{i=0}\sum_{j=0} C_{ij}\right)^{-1} \tag{2.9a}$$

$$\text{Sensitivity} = C_{11}\left(\sum_{j=0} C_{1j}\right)^{-1} \tag{2.9b}$$

$$\text{Specificity} = C_{00}\left(\sum_{j=0} C_{0j}\right)^{-1} \tag{2.9c}$$

$$\text{Precision} = C_{11}\left(\sum_{i=0} C_{i1}\right)^{-1} \tag{2.9d}$$

The networks as described in Table 2.3 were then compared against the validation set in the same manner where the $\varsigma(x), \tanh(x)$ architecture performed the best with 94.9%

accuracy, 84.8% sensitivity, 100% specificity and 100% precision. The results of Table 2.3 imply that the basic setup of all the networks tend to give false negatives as opposed to false positives which means good data can sometimes be rejected but bad data will always be identified. This is most likely due to the high acceptance threshold of 0.7, however a more canonical BCCE threshold of 0.5 gave rise to more false positives – a much more serious issue. More training data or a more complex NN architecture might resolve this issue at 0.7 acceptance tolerance.

With this current architecture, the softmax function did not produce effective learning ($> 95\%$ accuracy and $> 85\%$ in other categories) due to eventual gradient explosion caused by some of the $\partial_i \mathrm{s}(\mathbf{x})$ differentials during back-propagation. In fact, at best sensitivity, it predicted every training profile as 'bad' – not desirable for signal processing! This is most likely caused by the condensation of the 32 inputs overfitting to the output layer (Caruana et al., 2000) due to the open positive ranges of both the hidden and output layers. More sensitive CNN architectures were explored in the hopes of avoiding overfitting and to ensure that $e^{\mathbf{x}} \sim \sum e^{x_j}$ (see Table 2.1). These architectures were $(2048, 256, 32, 16, 32, 2)$, $(2048, 128, 64, 16, 32, 2)$ and $(2048, 128, 32, 16, 32, 2)$, with the central nodes being activated by the ReLU function. These were run for $20,000$ iterations on the same training set as before at a threshold of 0.7 and learning rate of 0.0004. Of these, the third arcitecture showed the most improvement for both the R and ς input activation functions with, respectively, 90.4/92.3% accuracy, 71.2/76.8% sensitivity, 100% specificity and 100% precision. This is still an undesirable level of error after this amount of training and, although the precision and specificity imply that no unwanted noise will enter the final TOA calculations, the relatively low sensitivity implies that a lot of good data is still being rejected.

The first four CNN architectures in Table 2.3 were arranged in a 5–layer node structure identical to the third architecture in the paragraph above. An identical training set to the 430 MHz set was made using L-band data, again using data from MJD $58402 - 58408$, with a learning rate of 0.0004 and threshold 0.7. These sets were then tested against L-band data for PSR J1829+2456 taken on MJD 58235 and used to excise potential RFI from that observing day. This day in particular was chosen because the uncleaned integrated pulse profile shows incredible baseline variation that leads to TOAs with high uncertainty that may not be very accurate. A crude S/N of each integrated pulse profile was calculated as the maximum signal in the profile divided by the off-pulse RMS and the S/N of the non-excised control profile. This allows for the direct quantifiable diagnostic of the effectiveness of each architecture; results are shown in Table 2.4. The profiles were also manually inspected (by eye) to ensure the individual outputs led to pulse profiles suitable for timing. These integrated profiles are shown in Figure 2.2.

Overall, the CNN architecture best suited for RFI excision was the $\varsigma(x), \sigma(x)$ setup with node structure: $(|\mathbf{X}|, 128, 32, 16, 32, 2)$. Although more advanced architectures have been designed for pulsar analysis (e.g. Zhu et al., 2014, who deal with the far more voluminous search-mode data so the number of neurons in a layer ~ 9000), these often require the use of a server or cluster whereas the setup explored above is completed in a reasonable time on a home computer.

Table 2.4: The signal-to-noise ratios for the integrated profile of PSR J1829+2456, observed by the Arecibo telescope on 58235 at a center frequency of 1400 MHz, after RFI excision based on four types of CNN architectures. These are compared with the raw 'uncleaned' integrated pulse profile. The profiles were dedispersed using the DM given in Table 3.2 and excised of RFI using the 5–layer NN activation architectures given below.

Architecture	S/N
No excision method	1.00
$\mathrm{R}(x), \sigma(x)$	1.63
$\varsigma(x), \sigma(x)$	1.90
$\mathrm{R}(x), \tanh(x)$	1.52
$\varsigma(x), \tanh(x)$	1.30

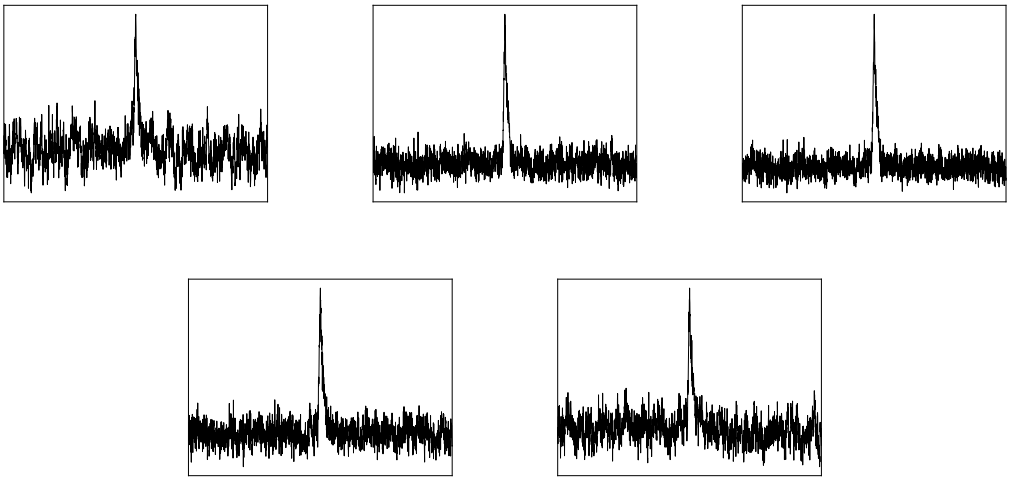


Figure 2.2: The integrated pulse profiles for each neural network architecture explored in Table 2.4. From top left to bottom right: No RFI excision; $\mathrm{R}(x), \sigma(x)$; $\varsigma(x), \sigma(x)$; $\mathrm{R}(x), \tanh(x)$ and $\varsigma(x), \tanh(x)$. Each x-axis represents one full phase rotation and the y-axes are all in arbitrary flux density units.

As mentioned, the CNN training administered here does not use pre-defined features but instead determines the best performance by comparison to the four metrics described throughout the section. All CNNs start off equal, with randomly assigned weights and then the output of successive training 'guides' the NN in the right direction, akin to a black box. As such, every pulsar put through this pipeline needs to undergo an initial training process. Weights and biases from one pulsar do not generally translate to other pulsars. However, effective training can be reached with $\sim$ 100 profiles so this is not usually an issue.

Analysis of the Vapnik–Chervonenkis dimension

As a final test of the overall effectiveness of the aforementioned NNs, a measure of the NN classifier's 'capacity' was made. Capacity is a blanket term for a network's complexity: how it learns, thinks and potentially overfits. The criterion used to measure the capacity was the Vapnik–Chervonenkis (VC) dimension for probabilistic classifiers (Vapnik and Chervonenkis, 1971); however it works well here due to the 'straight-line' threshold cut-off at the output layer. The VC dimension is formally the cardinality of the largest set that can perfectly divide (or *shatter*) an input set. For nodes, V, and node connections, E, Shalev-Shwartz and Ben-David (2014) give the scaling bounds of the VC dimension, D, for a sigmoid-like activated NN with initially random weights to be:

$$\Omega(|E|^2) \leq D \leq O(|E|^2 \cdot |V|^2) \tag{2.10}$$

where the $|...|$ means 'number of' and Ω and O are the lower and upper asymptotic functionals respectively[14]. Essentially, this means that a CNN with a high number of nodes in the j^{th} layer and a comparatively low number of nodes in the $(j+1)^{\text{st}}$ layer (i.e. high V, low E) will asyptotically have a constant VC dimension with decreasing V. On the other hand, a CNN with a large number of connections, regardless of node structure, will increase D with the square of the increase in connections. It is favorable to have a high VC dimension at each layer but not exceeding the number of training examples (in this case, 375). If $D > N$, overfitting may occur because there are more connections than examples so the network cannot be entirely resolved during back-propagation. The two layers which display the most issue at input are the input layer ($D \sim \Omega(|E|^2)$) and the final hidden layer ($D \sim \log_2(|E|)$) which could potentially explain the softmax function's inability to create an effective learner. The VC dimension is therefore also a measure of NN stability.

Further work into tweaking the CNN architecture could be done to increase all the metrics towards 100% however the NNs described here would not give rise to any difference in parameter measurements for the pulsars in this work. Given the small difference between the results shown here and a 100% efficient CNN, TOA precision is predicted to increase by only a few nanoseconds: far below the sensitivity of the telescope, so these results are as physically significant as the 'perfect' CNN.

[14]For Ω, the Knuth definition has been used (Knuth, 1976).

2.2.3 Excision via inverse phase–space analysis: a theoretical model

So far, RFI methods as explored in this thesis have employed statistically motivated methods for data excision. This means that each profile is observed as a whole and compared with the standard profile either by statistically determining RMS cut-off points (Section 2.2.1) and by determining the best cost function and applying the best case scenario (Section 2.2.2). The following section outlines a theoretical process for potential non-statistically motivated RFI excision to take place.

Assume that the standard profile of a set of pulsar observations can be expressed as approximately the sum of a number of Gaussian functions with general form:

$$g(t) = ae^{-\frac{(x-b)^2}{c^2}} \tag{2.11}$$

where a is the maximum flux density of that Gaussian, b is the mean and c is the standard deviation. The continuous Fourier transform of an unnormalised Gaussian function, such as is described above, is given by (making use of Eqs. 7.4.2 and 7.4.3 in Abramowitz and Stegun, 1972, p. 302):

$$\begin{aligned}
\tilde{g}(k) &= \int_{-\infty}^{\infty} g(t)e^{-\mathrm{i}kt}dt && (2.12)\\
&= \int_{-\infty}^{\infty} ae^{-\frac{(x-b)^2}{c^2}}e^{-\mathrm{i}kt}dt && (2.13)\\
&= a\int_{-\infty}^{\infty} e^{-\frac{x^2}{c^2}}e^{-\mathrm{i}(t+b)k}dt && (2.14)\\
&= ac^2e^{-\mathrm{i}bk}\sqrt{\pi}e^{-\frac{c^2k^2}{4}} && (2.15)
\end{aligned}$$

In other words, the Fourier transform of a Gaussian is itself a Gaussian – this is true for the discrete case as well. It follows that this is true for Gaussian summation, meaning that the Fourier transform of the standard profile with arbitrarily many components is a Gaussian with a set of overtones. It should therefore be possible to compare the Fourier transform of the standard profile, $\tilde{s}(k)$, with the Fourier transform of an individual profile, $\tilde{p}(k)$, in order to determine whether the profile contains a significant degree of RFI. This can be achieved by baseline removal with baseline $\tilde{s}(k)$. Profiles that follow roughly the same wave pattern as the standard profile, after baseline removal, will be on the order of a null profile whereas profiles with considerable RFI will display different fundamental frequencies when compared with $\tilde{s}(k)$. This is akin to fast-Fourier transform (FFT) RFI methods for pulsar search data (e.g. Ransom, 2001), however this theory takes the limit as the step size, $dt \to 0$, in essence removing the discrete aspect of the calculation. For non-zero dt, there is a resolution 'trade-off' in which nuanced signal information is lost and smoothed, however designing the FFT step size to be significantly less than an individual profile bin size becomes computationally expensive very quickly. FFT works well with highly factorisable data sizes (powers of 2) and, at best, can achieve a cost of $O(N \log N)$. The theoretical possibility of such a model is within the realms of quantum computing

however the FFT precision needed to satisfy the goals of this investigation remain within the realms of the discrete Fourier transform. The continuous aspect of this investigation comes from the Gaussian fitting procedure which returns function parameters as opposed to a set of output points (see Appendix B.1). Since these function parameters apply to continuous space, one *can* take the continuous form of the Fourier transform as valid for any initial bin size so long as the resolution of the observation was fine enough to obtain the desired features in the fit.

At the present time, we are limited in how to implement this method and it is not explored further expect for one comparison later on. However, as computing power gets better, we can hope to see developments in the continuous Fourier transform regime. One might argue that this method interpolates data that may not be there. This may be true, but it is *not* a smoothing of the data. Every data point is still present but added in with intermediate values. It is the hope of this section that others in years to come may attempt something similar on a much more massive scale.

2.2.4 Discussion of the models

A wide variety of methods have been explored to deal with the seemingly permanent problem of RFI in radio pulsar data, and the methods in this work by no means cover the entire breadth of study into noise handling. Since data for this thesis was taken in both incoherent search-mode and coherent fold-mode, the RFI excision methods investigated take into account both types of dataset. Search-mode data already has a number of RFI excision techniques that can be applied to it, such as pre-dedispersion zero-DM filtering (Eatough et al., 2009) and fast RMS σ–clipping via PRESTO (Ransom, 2001). Fold-mode RFI excision is a post-folding process, which means that the quality of the initial solution can greatly affect the pulse shape and therefore the ability of RFI excision software to accurately determine RFI-riddled channels.

In order to determine the optimum method of those discussed in this section, a set of TOAs were generated using PALFA search-mode data for PSR J1936+1805 spanning MJD $57924-58895$, refolded using the most up-to-date model ephemeris. These were then fit with TEMPO2 using the DE438 Solar System ephemeris. For the preprocessing, each set was excised of RFI using each technique as described in the above sections. All other variables (e.g. calibration factors and number of TOAs created per observation) were kept consistent between models, although not all excision techniques produced the same number of fittable TOAs. This is because TOAs below 1100 MHz were omitted due to the receiver cut-off for ALFA, and because the maximum allowed TOA uncertainty was set to 100 μs to account for the small number (< 10 for the entire data span) of TOAs above 1100 MHz, for which no method could excise RFI from effectively. In each case, the reduced χ^2 and the residual RMS, $T_{\rm res}$, were monitored with the number of fit parameters remaining as five throughout. The five fit parameters were the pulse frequency F0, the pulse frequency derivative F1, the right ascension, the declination, and the dispersion measure. The results are outlined in Table 2.5 and the residuals are plotted in Figure 2.3. The training sets for the CNNs were made using data from MJD $58330-58359$, with null profiles amounting

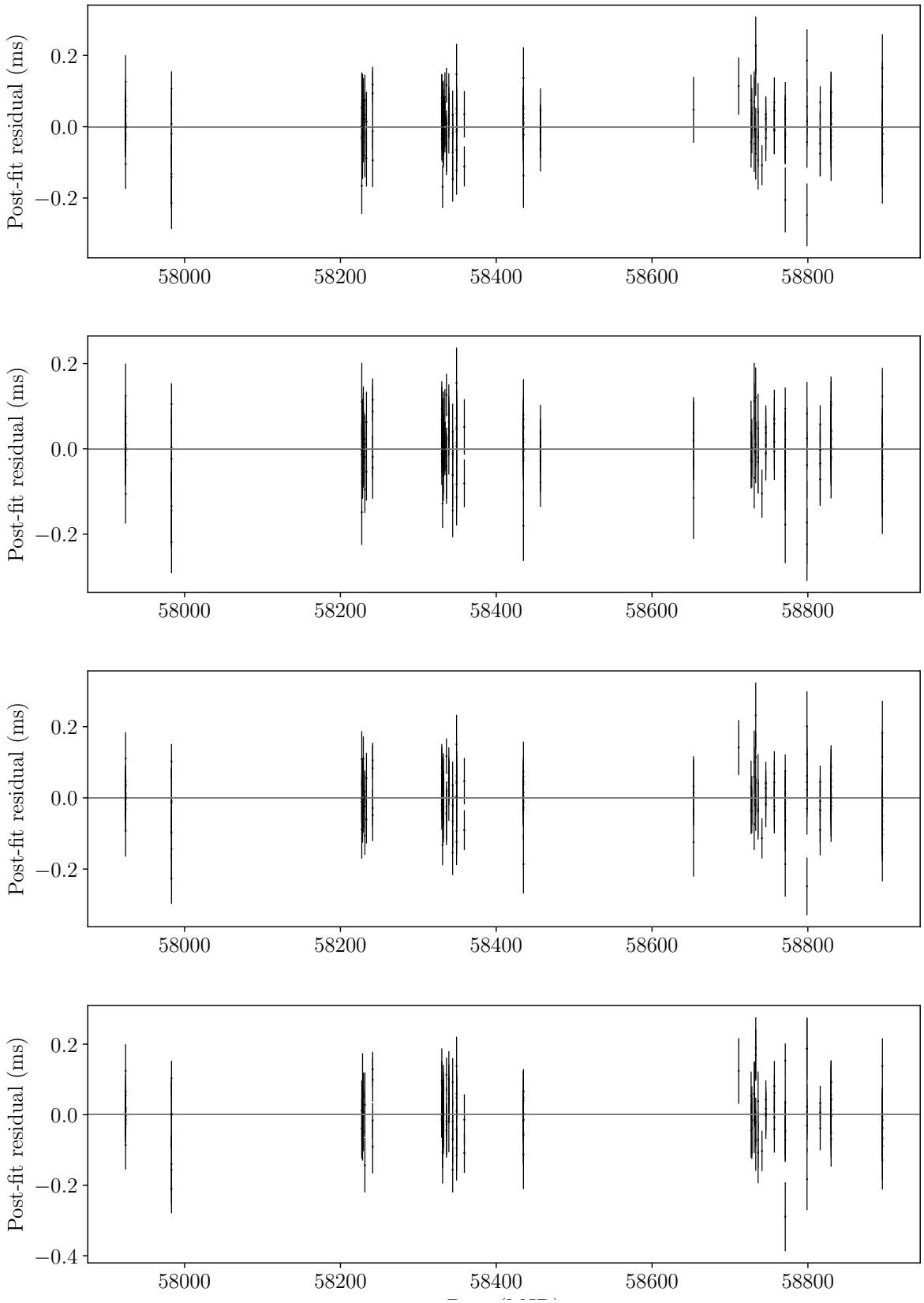


Figure 2.3: Post-fit timing residuals representing the results of using four RFI excision methods for the isolated pulsar PSR J1936+1805. From top to bottom: Chauvenet's criterion ($\sigma = 2$); $\varsigma(x), \sigma(x)$ CNN; $\varsigma(x), \tanh(x)$ CNN and Fast Fourier transform ($\delta t = 1$ phase bin).

to 5% of the number of true profiles in the training sets.

Table 2.5: A comparison of four RFI excision algorithms for use on refolded data from PSR J1936+1805 observed by AO between MJD 57924 − 58895. The threshold for Chauvenet was set to 2σ and the NN was given the same 5–layer node structure as shown in Table 2.4. The TOAs were fit using TEMPO2 using the DE438 Solar System ephemeris. In each instance, five timing parameters were fit (see text for specifics).

Model	# TOAs	$\chi^2_{\rm r}$	$T_{\rm res}$ (μs)
Chauvenet	158	1.171	68.761
$\mathrm{CNN}_{\varsigma,\sigma}$	159	1.056	64.257
$\mathrm{CNN}_{\varsigma,\mathrm{t}}$	156	1.137	67.400
FFT	138	1.263	73.039

In the above analysis, FFT performed the worst at accurately determining which profiles to omit from the averaging process. Upon inspection of each averaged profile responsible for each TOA in the FFT TOA set, it was clear that the FFT method fails to identify subtle cases of RFI. This is due to the limiting resolution that discrete Fourier transform allows when compared with a continuous Fourier transform; this, alongside its long computation time, make it the least desirable option. As expected, the $\varsigma(x), \sigma(x)$ CNN performed best and was the basis for the fold-mode RFI excision used in Chapter 4.

2.3 Flux calibration

Another facet of downstream data analysis for pulsars is that of flux calibration. The orthogonal dipole feeds at the telescope receiver accept signals that are propagated through separated channels in the IFLO chain, subject to e.g. amplification, filtering, etc.

Flux calibration will scale each polarisation channel seperately based on a reference source of known polarisation and flux density thereby allowing for any discrepancy between the received signal from each feed to be accounted for. This process also allows for the conversion of flux units from arbitrary instrumental signal units (hereon referred to as 'counts') to more conventional units, such as Jy. The flux density recorded at a particular feed is a function of the specific telescope aperture temperature efficiency, known as the telescope *gain*, ($G_{\rm tele}$, where 'tele' is replaced by the relevant telescope name), and the polarisability of the night sky at the time of observation. Calculating an accurate value for the signal flux density ensures the proper scaling of the orthogonal polarisation components post–observation, with the end goal of recreating the intrinsic pulse profile. With much of the same motivation as for RFI mitigation (see Section 2.2), flux calibration therefore attempts to correct for any sources of flux density error. It should be noted that flux calibration is often not necessary for pulsar timing as TOAs are calculated as horizontal offsets from the pulse peak, however flux calibration was administered in this work in order to fully characterise each pulsar's emission so that they are ready for any further analyses.

When performing flux calibrations on radio sources, one can compare the pulsar signal against that of a source of known flux density and stable spectral index such as a quasar or radio galaxy. Unpolarised standard calibrators are often preferred so that it can be assumed

that the flux density is equal for both hands of polarisation. There are two standard ways to record flux densities of standard calibrators. The first method requires information of the continuum source's flux density, $S_{\rm c}$ in Jy, at a given frequency, ν_0 in MHz as well as its spectral index, α, which is assumed to be constant in the pulsar's observation frequency range. To calibrate the flux density to the observing frequency, $\nu_{\rm obs}$, the transformation is given by

$$S_{\rm c}\left(\nu_{\rm obs}\right) = S_{\rm c}\left(\nu_0\right)\left(\frac{\nu_{\rm obs}}{\nu_0}\right)^{\alpha} \tag{2.16}$$

The second way of calibrating flux contains a list of polynomial coefficients $\{a_0, a_1, ...\}$, satisfying

$$\log_{10} S_{\rm c} = \sum_{k=0}^{n} a_k \log_{10}^{k}\left(\nu_G\right) \tag{2.17}$$

where ν_G is the centre frequency in GHz. Most calibrations that use Equation 2.17, such as the calibrations made in this thesis, are quoted at the $n = 3$ level as this level gives flux densities to within $\sim 1\%$ and adding a fourth polynomial parameter often does not show an appreciable difference in the fit (see Baars et al., 1977, for a more engaging account of polynomial flux calibration.). Continuum source data for this thesis was provided by the North American Nanohertz Observatory for Gravitational Waves (NANOGrav) and the pulsar data files were matched to the closest available calibrator observation.

Flux calibration as performed in this thesis also uses a phase-correlated pulsed signal from a noise diode, with a stable oscillating frequency $\nu_{\rm ND} = 25$ Hz, injected at the receiver. This is much more accurate than calibration via S/N or from the system noise as it does not assume that the contribution to the system noise temperature, $T_{\rm sys}$ $(= G_{\rm tele} S_{\rm sys})$, is the same for both the calibration source and the pulsar source. This is often not the case and cannot be ignored, even approximately. The pulsar was observed for a small amount of time (typically 90 seconds) with the noise diode switched 'on', and then observed again for an equal duration with the noise source switched 'off'. This allows the contribution of the noise source to the observed flux density to be calculated. These calibration profiles were folded at a period of $1/\nu_{\rm ND} = 0.04\,{\rm s}$, resulting in a square-wave profile that switches between 'on' and 'off' signal levels at a phase of 0.5. The characteristic temperature, $T_{\rm cal}$, due to the diode is then given as a function of the average flux counts, K (for 'kount'), by:

$$T_{\rm cal} = T_{\rm sys}\left(\frac{K_{\rm off}^{(25{\rm Hz})} - K_{\rm off}^{(0{\rm Hz})}}{K_{\rm off}^{(0{\rm Hz})}}\right) \tag{2.18}$$

where the superscripts on K denote the noise diode frequency with 0 Hz corresponding to the non-diode state. The K subscripts 'off' and 'on' (below) denote whether the telescope is viewing the continuum source or not. Similarly, the characteristic temperature due to the radio signal, $T_{\rm src}$, is given by:

$$T_{\rm src} = T_{\rm sys}\left(\frac{K_{\rm on}^{(0{\rm Hz})} - K_{\rm off}^{(0{\rm Hz})}}{K_{\rm off}^{(0{\rm Hz})}}\right) \tag{2.19}$$

The calibration flux density is found by combining Equations 2.18 and 2.19 and dividing the resultant $T_{\rm cal}$ expression by the telescope gain, which is $11\,{\rm K\,Jy^{-1}}$ at AO for both observing frequencies[15]. After mathematical manipulation, $S_{\rm cal}$ is therefore given by:

$$S_{\rm cal} = S_{\rm src}\left(\frac{K_{\rm off}^{(25{\rm Hz})} - K_{\rm off}^{(0{\rm Hz})}}{K_{\rm on}^{(0{\rm Hz})} - K_{\rm off}^{(0{\rm Hz})}}\right) \tag{2.20}$$

where S is in Jy. The qualitative relationship between Equations 2.18, 2.19 and 2.20 are shown in Figure 2.4, which has been created from real calibration data although the time (x) axis is not to scale.

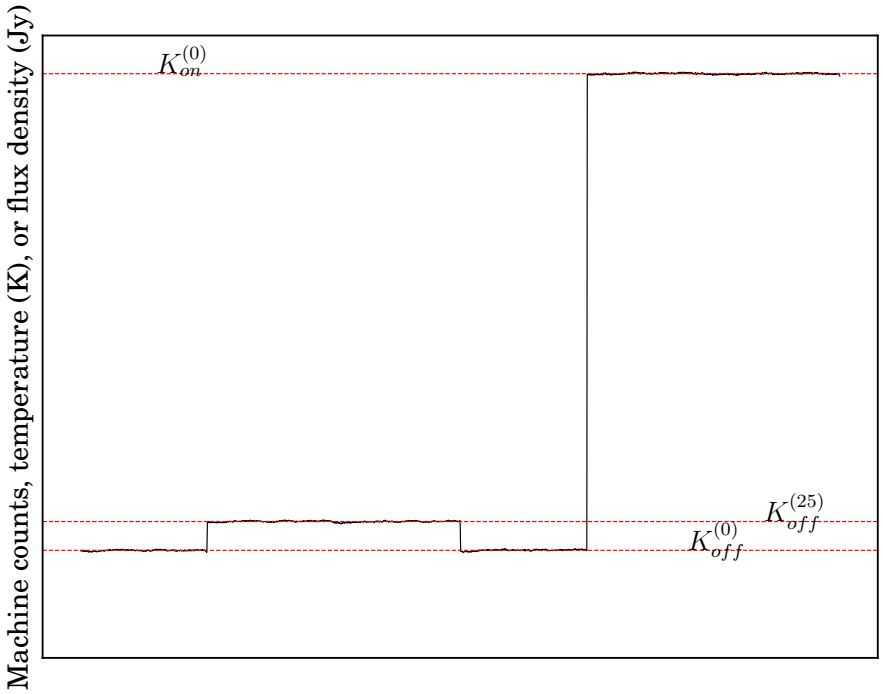


Figure 2.4: Schematic of the continuum source flux density contributions for QSO B1442 (J1445+0958). This observation was taken on MJD 58067 at a center frequency of 1380 MHz with an oscillating noise diode set to 25 Hz. The profile seen here is that of the first polarisation channel.

To obtain the calibration factor, η, in ${\rm Jy\,count^{-1}}$, $S_{\rm cal}$ is divided by the flux density contribution of the noise diode when the telescope is pointing at the pulsar. This aids to account for effects arising from e.g. variable telescope elevations and atmospheric conditions if necessary. This is given by:

$$\eta = \frac{S_{\rm cal}}{K_{\text{on-psr}}^{(25{\rm Hz})} - K_{\text{on-psr}}^{(0{\rm Hz})}}{\rm Jy\ count^{-1}} \tag{2.21}$$

[15]The general system parameters for all of the Arecibo receivers can be found at http://www.naic.edu/~astro/RXstatus where the value of 11 for the gain was obtained. More specific measurements for gain can be obtained from each receiver's homepage but the variance in gain is insignificant compared to the timing uncertainties in this work.

This has the effect of normalising the intensity between each hand of polarisation (see Figure 1.3) before obtaining the total intensity power profile. The relevant calibration factor was calculated for all frequency channels, which were then interpolated via 1D cubic-spline fitting using SciPy (Virtanen et al., 2020). The interpolated functions were applied to each polarisation channel of the data giving the flux density conversion factors. The resulting factors were applied to the data (along with the RFI mask) before the polarisation components were combined to obtain a total power signal.

In the rare cases where observations were conducted $\gtrsim 30\,\mathrm{d}$ from the nearest continuum source observation and therefore no value for T_{cal} could be calculated, the calibration factor was taken to be the reciprocal of the off-pulse RMS for each polarisation channel. This also has the effect of essentially dividing the profile by the S/N to ensure a more accurate pulse profile.

To further develop the calibration as described here for each polarisation channel, one can use the more computationally expensive 'polarisation calibration' method. This method transforms each Stokes vector independently via a Mueller matrix (Mueller, 1948). Until the 21$^{\mathrm{st}}$ century, computing power was insufficient to perform the back-and-forth matrix inversions necessary for polarisation calibration, however this is no longer the case. Polarisation calibration is arguably more useful for highly polarised pulsars and is not needed in this work as the vast majority ($> 95\%$) of signal, with respect to the four Stokes vectors, lies in the intensity (I) for all five pulsars discussed and so only intensity was used to make TOAs. Polarisation calibration was administered before flux calibration for all of the fold mode timing data used in this work using the `pac` command in the PSRchive software suite (Hotan et al., 2004), specifically using the SingleAxis model which assumes that the polarisations of the receivers are orthogonal.

Chapter 3

PSR J1829+2456: a highly relativistic neutron star binary

More than 2600 neutron stars presenting as radio pulsars have been found since their first discovery. This large population has given a wealth of physical and universe-wide tests through continued pulsar timing (see Section 1.2). Timing studies have resulted in: robust models for the free electron density in the Local region (Cordes and Lazio, 2002; Cordes and Lazio, 2003; Yao et al., 2017), which can be used to estimate pulsar distances from the dispersion measure; a beam model for the EM jet (Dyks, 2017); and the improvement of solar system ephemerides by way of fitting the planet masses to precise timing results (Champion et al., 2010). Despite our collective effort over the past 50 years, the physics and chemistry within and surrounding neutron stars is still poorly understood due to their highly exotic nature and cataclysmic birth mechanisms. There is added complexity in these stellar evolution models when the pulsar exists in a binary system – mass is accreted either stably or unstably, ultimately depending on the timescale of interaction (see Section 1.3).

Since neutron stars in binary systems are very gravitationally active, it is reasonable to suggest that obtaining precise mass measurements for NSs will lead to a more complete understanding of the evolution of these types of stars as well as provide tests for theories of gravity. In GR, gravity is modelled entirely by the spacetime metric however in the strong-field regime, it is possible, for instance, that non-linear scalarisation effects can take place due to the NS matter (Damour and Esposito-Farèse, 1993, the so-called 'DEF' model, which is energetically more favourable than GR) or spontaneous scalarisation due to the extreme space-time curvature in these regions (Silva et al., 2018, the so-called 'Gauss-Bonnet Coupling'). These *scalar-tensor* theories of gravity satisfy the Einstein field equations in the weak field (in which any scalar field is essentially 0) but are, at present, not well constrained in the strong-field. This is due to wide uncertainties on DNS masses and the many possible choices for NS equation-of-state (see Shao et al., 2017, for an overview of why binary pulsar timing can aid in our understanding of these perturbations).

Much of the work detailed in this chapter has been published as (Haniewicz et al., 2021).

Binary systems can be observed in order to find NS masses through a modelled theory of gravity (such as GR) however only about 4% of all known pulsars exist in binaries, with only 21 such systems known to be double neutron stars. The evolution of DNSs is thought to be a highly selective set of events, which is echoed in the observed tight constraints of astrometric and orbital parameters for these systems.

Presented here are the findings from continued studies on the double neutron star system containing PSR J1829+2456 (Champion et al., 2004; Champion et al., 2005) which include a precise measurement of the component masses as well as more precise observations of the dispersion measure, eccentricity and proper motion. These values are compared against the DNS population in order to better constrain the current evolution models.

3.1 Mass and velocity distribution among Galactic DNS systems

The mass distribution of DNSs has been studied by many groups. Of the only 21 known DNS systems, 19 are Galactic and only a handful of these have precise mass measurements. This is unfortunate as the evolution and supernova mechanism leading to these systems is closely tied to their current masses (Pejcha et al., 2012). The first widely accepted comprehensive mass distribution study (Thorsett and Chakrabarty, 1999) found a NS mass distribution of $1.38^{+0.1}_{-0.06}\,M_\odot$ to the 1σ level. This distribution was fit to a Gaussian (Özel et al., 2012) which was later updated (Özel and Freire, 2016) and found the most likely values of the mass to the 1σ level to be $(1.33 \pm 0.09)\,M_\odot$, however this is likely skewed slightly by the low population size. The NS mass distribution for all binary pulsar systems (i.e. including NS–WD systems) was found to be $1.54 \pm 0.23\,M_\odot$ for recycled NSs and $1.49 \pm 0.19\,M_\odot$ for slow pulsars, although the validity of the mean in these latter fits is contested by a two-peak distribution model (for a compelling argument of the distribution bimodality, see Antoniadis et al., 2016; Farrow et al., 2019). In any case, the distribution of NS masses is far narrower in DNS systems, implying that the conditions required to sustain a DNS system after supernova are tightly constrained.

Recently, two DNSs with highly asymmetric NS masses have been discovered (Martinez et al., 2015; Ferdman et al., 2020). This has not only expanded the range of NS masses observed in DNSs, but showed the existence of asymmetric DNS systems. The latter system is expected to merge within 470 Myr which suggests that the population of such merging asymmetric DNSs might be substantial, occupying about 10% of the known merging DNS population. Establishing more firmly the size of this population will be of particular importance for the interpretation of DNS mergers in future data from LIGO, Virgo and Kagra.

3D space velocities are difficult to estimate due to the way velocity is determined: through the precise measurement of proper motion components. This 2D motion is observed as a secular change in the pulsar's position in the sky over a number of observations and is observed as a 'smearing' effect in the pulsar profile if the positions are left unchanged. Tangential (or transverse) velocities can be calculated directly from the proper

motion measurement and the distance, which can be best estimated to 20% (Yao et al., 2017) using interstellar dispersion, or on the order of the parallax error if known; radial velocities along our line of sight are far harder to accurately assess. Taking the extreme case where the pulsar is moving away from Earth in only a radial direction would result in a proper motion measurement of 0 – the pulsar appears not to move at all! So calculation of the radial velocity is highly dependent on measurement of the tangential velocity (Helfand and Tademaru, 1977). NSs in Galactic DNSs are mostly observed to have low tangential velocities ($\sim 10\,\mathrm{km\,s^{-1}}$), although some, such as the Hulse–Taylor binary (PSR B1913+16, $v_{\mathrm{trans}} \approx 150\,\mathrm{km\,s^{-1}}$; Hulse and Taylor, 1975) and PSR B1534+12 ($v_{\mathrm{trans}} \approx 120\,\mathrm{km\,s^{-1}}$; Fonseca et al., 2014) have been shown to have much higher space velocities ($\sim 100\,\mathrm{km\,s^{-1}}$). The implications of this will be discussed further in Section 3.6.

3.2 Observations

PSR J1829+2456 is a relatively old ($\tau_c \approx 13$Myr) recycled pulsar with a rotational period of 41.0 ms, found in the constellation Hercules. It was initially discovered from data taken during a 1999 drift-scan survey using the 430 MHz Gregorian dome receiver system at the Arecibo Observatory (AO). At the time of its discovery, the dispersion measure (DM) was found to be $13.9\,\mathrm{pc\,cm^{-3}}$, which estimated a distance of 1.2 ± 0.36 kpc to the system using the NE2001 Galactic ionised electron distribution model (Cordes and Lazio, 2003). This distance is likely overestimated due to the dispersive effects of the Gould Belt, a dense 3000 ly span of O and B type stars that eclipses our line of sight to the system (Gehrels et al., 2000; Grenier, 2000). The NE2001 does not take this region into full account and a more reliable estimated distance may come from using the YMW16 electron distribution model (Yao et al., 2017), which models corrections for a fourth spiral arm in the Milky Way and to the Local Bubble. The YMW16 gives the distance to be 0.91 ± 0.18 kpc, however, given that population studies on this region are still on-going, the estimate given by this model is likely a slight underestimate (see Kerr et al., 2019, for some recent Gould Belt developments). Given this distance estimate, it is possible that the distance may be measureable using Very Long Baseline Interferometry (VLBI), but this has not been attempted since the pulsar's discovery.

Upon initial timing, the advance of periastron, $\dot{\omega}$, was the only significant post-Keplerian parameter measured, found to be $0.2919 \pm 0.0016\,\mathrm{deg\,yr^{-1}}$. This measurement determined the total binary mass to be $2.59 \pm 0.02\,M_\odot$. As component masses can only be disentangled with a measurement of two or more PK parameters, only limiting values of $m_{\mathrm{p}} < 1.38\,M_\odot$ and $1.22\,M_\odot < m_{\mathrm{c}} < 1.38\,M_\odot$ could be placed on the system for the pulsar and companion masses respectively, making use of the mass function value reported to be $0.2942\,M_\odot$. These mass boundaries do not conclusively determine the companion to be a NS as the maximum companion mass is well below the Chandrasekhar Mass of $1.4\,M_\odot$ (Mazzali et al., 2007), however the moderate orbital eccentricity along with the mass bounds implies a DNS for reasons outlined in Sections 1.3 and 3.6.

Initial timing observations of PSR J1829+2456 began on MJD 52785 (26 May 2003) using the Penn State Pulsar Machine (PSPM) at a centre frequency of 430 MHz, and the

Wideband Arecibo Pulsar Processor (WAPP) centerd at 1400 MHz at Arecibo. Several observations were carried out using the Green Bank telescope (GBT) at 350 MHz in August 2006, but due to low signal-to-noise, only 10 time-of-arrival (TOA) measurements could be salvaged due to overly pervasive radio frequency interference in that data set. A full description of the former data set and its analysis can be found in (Champion et al., 2004) for the discovery and (Champion et al., 2005) for the follow-up.

The most recent observing campaign for PSR J1829+2456 ran since July 2017 until August 2020 after a nearly ten year hiatus in observations. Two dense campaigns of four epochs within one week were also conducted between 58402 – 58408 and 58748 – 58752 in order to provide better orbital phase sampling. All these observations were conducted at the 305-m Arecibo radio telescope roughly every 4 weeks using the Puerto Rico Ultimate Pulsar Processing Instrument (PUPPI) coherent de-dispersion backend. Preliminary calibration procedures were undertaken with $\nu_{\mathrm{ND}} = 25$ Hz over a 90 second calibration window (45s with the diode turned “on" and 45s with it turned “off").

Two frequency bands were used during the observations with centre frequencies of $\sim$ 1400 MHz, known as the L–band, and $\sim$ 430 MHz. These receivers had bandwidths of 800 MHz and 100 MHz, respectively, at the time of observing although it is noted that the L–band receiever cut-off is at around 1100 MHz. The data from each receiver was compartmentalised into 1.5625 MHz–wide channels, giving 512 total channels for the L–band observations and 64 channels for the 430.

A few physically significant things were taken into account to ensure the most precise measurements:

- The spin period, P, was updated for the current epoch of each observation based on the last known parameter of P using the Taylor expansion outlined in Section 1.2. An inaccurate spin period will not create a correct pulse profile when the time series is folded. Sometimes this is not an issue as pulsar spin–down rates are so slow when compared to gaps between observations, although spin–down rates are measurable and ought to be taken into account. Updating P using an expansion was necessary to do for PSR J1829+2456 because the ephemeris file was not updated after each observation, but rather every now and then. This, combined with the narrow pulse shape for PSR J1829+2456, requires P to be as accurately determined as possible.

- The data were coherently dedispersed using the procedure outlined in Section 1.2.3 to negate the sub–band frequency effects of the beam structure and the ISM to the pulse arrival time.

- Data from the 9 GPUs at Arecibo were combined and folded to create a series of profiles, varying in both frequency and time offset, to be used for timing.

A standard profile (Section 1.2) was created in an iterative manner, beginning with averaging all the folded data from a particular backend together and using the resulting profile’s total intensity as the template. RFI excision was conducted on data from the correct frequency band by cross-referencing against this profile using the method explored

in Section 2.2.1. This was done by fitting the current template to four Gaussian curves to obtain a smooth standard profile, allowing for a clear distinction between on-pulse and off-pulse regions. Individual profiles were then accepted or rejected based on a 95% Gaussian confidence in the distribution of off-pulse region root mean squared (RMS) values. After RFI excision, a new profile was made in the same way as described above by averaging the newly RFI excised data.

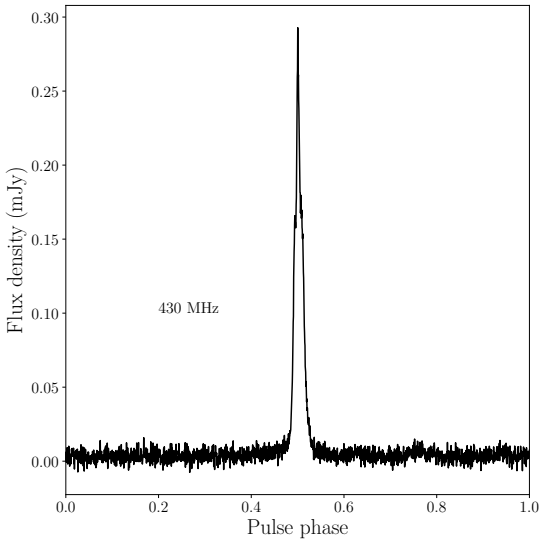


Figure 3.1: The integrated profile of PSR J1829+2456 at 430 MHz constructed from data taken between MJD 57950 and 58948.

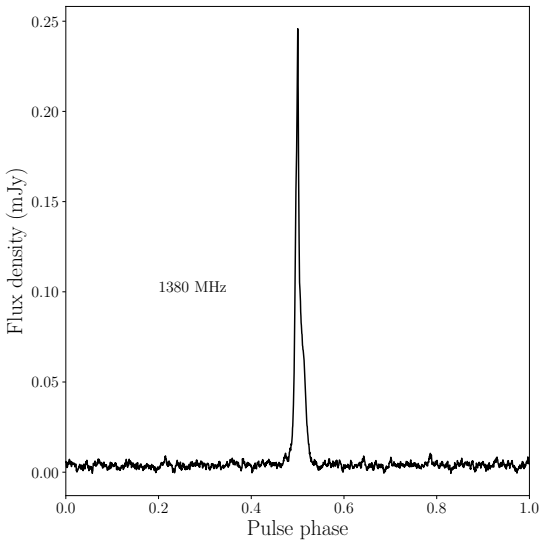


Figure 3.2: The integrated profile of PSR J1829+2456 at 1380 MHz (L–band) constructed from data taken between MJD 57950 and 58948.

The cleaned data were flux calibrated using the method outlined in Section 2.3 by comparing against observations of the stably polarised quasar QSO B1442 (J1445+0958) as a continuum source at the closest available dates to the PSR J1829+2456 observations; the largest time difference between the timing data and calibration data sets was nine days.

Data for this source was provided by the NANOGrav collaboration. After flux calibration, the fully processed data were once again used to create a standard profile for the band. Following this, an initial set of pulse times of arrival (TOAs) were generated, and a timing solution was fit to these TOAs. This was used to re-fold and phase-realign the PUPPI data.

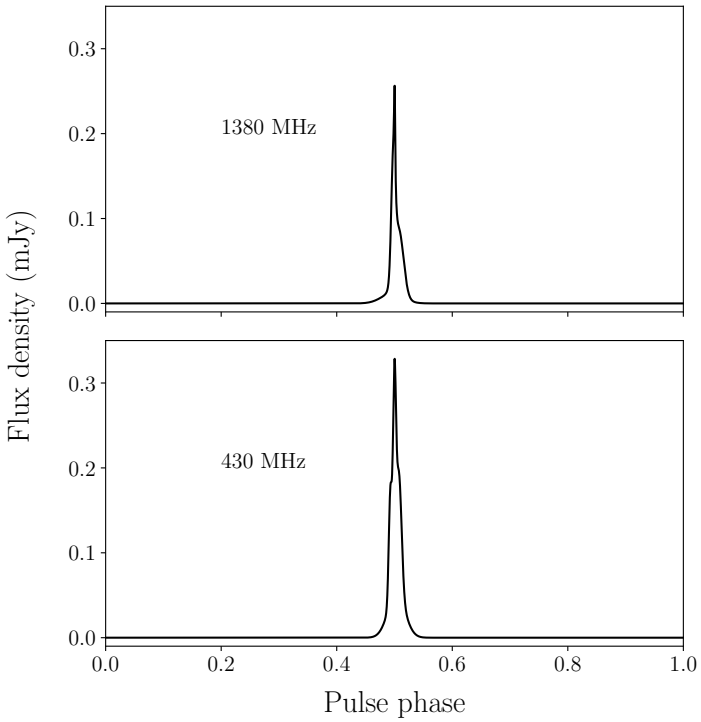


Figure 3.3: The noise-free standard profiles of PSR J1829+2456 constructed from all data taken with the PUPPI coherent dedispersion backend at Arecibo, up to and including MJD 58948. Top: L-band profile. Bottom: 430-MHz profile. Both standard profiles were fit with four Gaussians and are visually to scale relative to each other.

Using the same RFI excision masks and calibration factors as calculated in the previous step, the phase-aligned profiles were used to create the final standard profiles for each observing band, T_ν (where ν is the band centre frequency), by fitting the resulting points to another set of four Gaussian curves. This final step was done in order to ensure the template accurately reflected the intrinsic pulse shape, resulting in minimised timing residual errors. The raw integrated profiles for the 430 and the L–band data are given in Figures 3.1 and 3.2, respectively, and the Gaussian-smoothed standard profiles for both the 430 and the L–band are shown in Figure 3.3 and the breakdown of the Gaussians used to create them are detailed in Appendix B.1. All data manipulation was administered using the PSRVoid Python package (see Appendix C).

3.3 Timing analysis

In all, 1246 new pulse times-of-arrival were calculated from the data set by determining a phase offset for each resulting data profile through cross-correlation with the standard profile for each observing band. A careful analysis of the validity of TOAs created with this template so generally will be discussed further in Section 3.5. This phase shift was then converted to a time offset using the rotation period at the specific epoch corresponding to the individual data profile (Taylor, 1992). TOAs were created from time averaged sub-integrations of the complete Fourier transformed time series of about 9 minutes for the 430 MHz data, corresponding to 10 TOAs per observation, and about 3 minutes for the L-band data. The accurate time stamp of the pulse arrival time is given by the local H–maser correction in Puerto Rico with respect to UTC given by GPS satellite data. This is the time stamp that is compared against when cross-correlating the pulse. Details on pulsar timing and observation setup can be found in Section 1.2 and more specifically in Chapter 2. The L-band data were further divided into four frequency subbands centered at approximately 1680 MHz, 1480 MHz, 1280 MHz and 1080 MHz, however due to the aforementioned receiever cut-off, the TOAs created from the 1080 MHz subband were omitted from further analysis as well as TOAs with uncertainties larger than 25μs. In total, this resulted in 934 L–band and 314 430 MHz TOAs.

The TOAs were then appended to the 153 existing TOAs reported by Champion et al. (2005). These were fit within the TEMPO2 pulsar timing software package (Hobbs et al., 2006a; Edwards et al., 2006) using the JPL DE435 planetary emphemeris model (Folkner et al., 2016) and the TT(BIPM19) clock correction (Guinot, 1988) to calibrate to the Solar System barycenter (SSB) which is, to good approximation, an inertial reference frame. Where TT(BIPM19) could not be used (i.e. for the final three days of data), a correction was made in accordance with BIPM guidelines[1]. TEMPO2 fits all TOAs to an existing model ephemeris via a weighted least-squares fit, and outputs a set of timing residuals, which are the differences between the observed TOAs and those predicted from the current model. In all, 1399 TOAs were fit, spanning 17.1 years in total, at frequencies centred around 350 MHz, 430 MHz, 1280 MHz, 1480 MHz and 1680 MHz. TEMPO was occasionally used to verify parameter estimates, although not extensively.

Data taken from different telescopes and telescope frontends is essential to carrying out precise pulsar timing. In total, five different receiver frontends from two different telescopes have been used to obtain the current timing fit. The new PUPPI backend at the AO has provided this study with excellent data as have the older Arecibo hardwares. This, in combination with the TOAs made at the GBT, has given a good fit. One issue is that different telescopes and backends all have different operating systems as well as different engineering (as deep as them having used cables of differing resistivity). Most observatories use a different TOA calculation approach; for example, the coherent online folding at the AO provides a cleaner standard profile than the incoherent folding undertaken for the GBT data. As such, the UTC clock corrections to the SSB made above cannot fully account for all time delays. TEMPO2 has the option to fit for constant time offset "jumps", based

[1] ftp://ftp2.bipm.org/pub/tai/ttbipm/TTBIPM.2019

on the backend, in relation to a reference site. These are usually on the order of $\sim 100\mu$s, whereas typical TOA uncertainties are much less than this (on the order of the RMS). The jump fits for the observations were made against the PSPM–327 MHz TOAs for reference, although jumps fit to every other backend gave the same timing solution. All post–fit jumps were within the range $10 - 50\mu$s but are not reported. Theoretically, the WAPP data could be phase-connected with the PUPPI data however due to the time between each set of observations, determining phase connection is near-impossible so jumps were made between PUPPI and WAPP data. Residual errors were calculated from the propagation of uncertainty in the phase shift calculated in the cross-correlation process. A breakdown of each observation campaign is shown in Table 3.1.

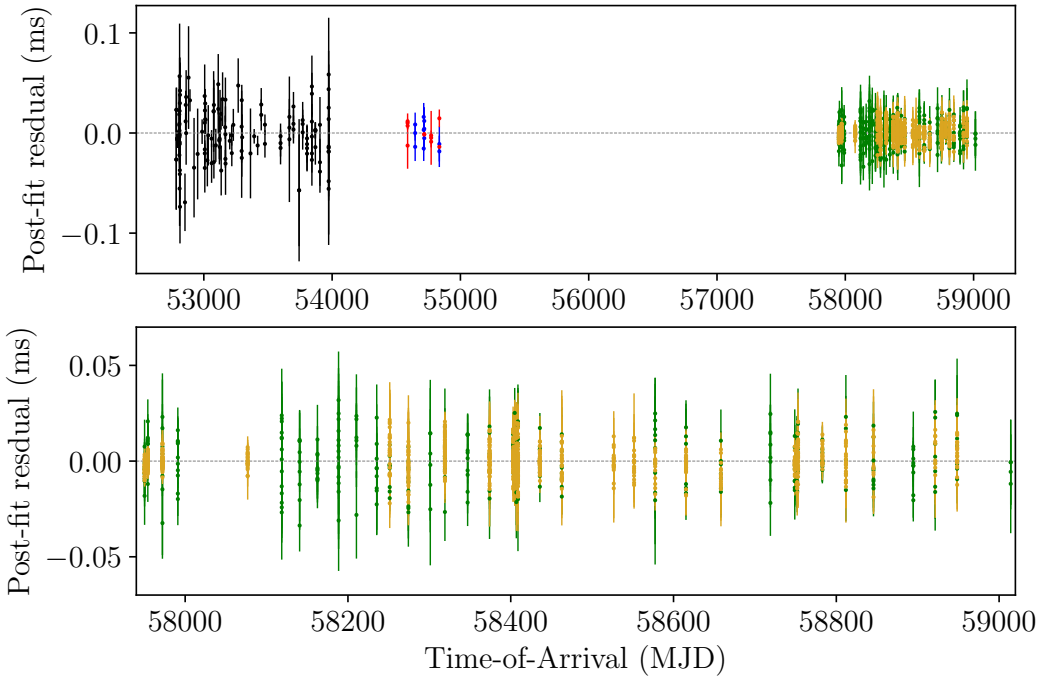


Figure 3.4: Post-fit residuals in milliseconds, as a function of TOA in MJD, for PSR J1829+2456 determined by the DDGR timing model. Top: all available Times-of-Arrival. Bottom: the new observations in green (430 MHz) and gold (1400 MHz).

Binary models

TEMPO2 offers a variety of binary fitting models with which to fit the current ephemeris to new TOAs. The Damour–Deruelle timing model (DD, Damour and Deruelle, 1985; Damour and Deruelle, 1986) allows for five individual post-Keplerian (PK) binary parameters to be measured in a theory-independent way using observation: the advance of periastron, $\dot{\omega}$, the orbital period decay, $\dot{P}_b$, the Einstein delay, γ, the Shapiro delay range parameter, r, and the Shapiro shape parameter, $s(\equiv \sin i)$. These PK parameters may then be re-cast under specific theories of gravity so that measuring two of these parameters can give direct individual mass measurements. The DDGR binary model (Taylor, 1987; Taylor

Table 3.1: Summary of time-of-arrival data for PSR J1829+2456.

Telescope	Instrument	Centre Frequency (MHz)	Bandwidth (MHz)	Span (MJD)	#TOAs	Weighted $\chi^{2\dagger}$	Weighted RMS† (μs)
Arecibo	PSPM	434.0	7.68	52785 – 53905	117	1.1873	17.0062
	PSPM	331.0	7.68	53027 – 53476	2	1.7937	19.9603
	WAPP	1378.6	100	54588 – 54835	11	2.0001	5.7050
	WAPP	319.6	100	54647 – 54835	11	0.9655	9.6925
	PUPPI	1384.4	800	57950 – 58948	934	1.1726	3.4843
	PUPPI	427.2	100	57950 – 59014	314	0.9938	10.7057
GBT	GASP	350.0	16	52972 – 52973	10	0.5231	32.5539
Overall		676.5	–	52785 – 59014	1399	1.1584	3.9640

†From the DDGR binary fit.

Table 3.2: Timing solution for PSR J1829+2456.

Fit and data-set		
Data span (yr)	17.1	
Date range (MJD)	52785.3 − 59015.3	
Number of TOAs	1399	
Solar System ephemeris	DE435	
Clock correction procedure	TT(BIPM19)	
Reference timing epoch (MJD)	55899.8	
Binary model	DDH	DDGR
RMS timing residual (μs)	3.967	3.964
Observed quantities		
Right ascension, $\alpha_{\rm J2000}$	$18^{\rm h}29^{\rm m}34\overset{\rm s}{.}66838(6)$	
Declination, $\delta_{\rm J2000}$	$24^\circ56'18\overset{''}{.}2007(12)$	
Rotation frequency, ν ($\rm s^{-1}$)	24.384401411044(6)	24.384401411040(6)
First derivative of rotation frequency, $\dot{\nu}$ ($\rm s^{-2}$)	$-2.9403(13)\times10^{-17}$	$-2.9395(14)\times10^{-17}$
Dispersion measure, DM ($\rm cm^{-3}pc$)	13.706(2)	13.707(2)
$\dot{\rm DM}$ ($\rm cm^{-3}$ pc $\rm yr^{-1}$)	−0.0019(4)	−0.0022(4)
$\ddot{\rm DM}$ ($\rm cm^{-3}$ pc $\rm yr^{-2}$)	0.00023(4)	0.00024(4)
$\dddot{\rm DM}$ ($\rm cm^{-3}$ pc $\rm yr^{-3}$)	$-7.0(1.6)\times10^{-6}$	$-7.1(1.6)\times10^{-6}$
Proper motion in right ascension, μ_α ($\rm mas\,yr^{-1}$)	−5.51(5)	−5.51(6)
Proper motion in declination, μ_δ ($\rm mas\,yr^{-1}$)	−7.75(7)	−7.82(8)
Binary period, $P_{\rm b}$ (d)	1.176027952868(11)	1.17602795281(15)
Orbital eccentricity, e	0.13914374(13)	0.13914387(11)
Projected semi-major axis of orbit, x (lt-s)	7.236845(2)	7.236844(5)
$\dot{x}$ (lt-s $\rm s^{-1}$)	$-2.3(5)\times10^{-14}$	
Longitude of periastron, ω (deg)	229.9353(2)	
Epoch of periastron, T_0 (MJD)	52848.5797762(7)	
Advance of periastron, $\dot{\omega}$ ($\rm deg\,yr^{-1}$)	0.293189(14)	–
Orbital period decay, $\dot{P}_{\rm b}$	$-2.9(1.2)\times10^{-14}$	–
Non–GR contribution to orbital decay, $\dot{P}_{\rm b}^{\rm X}$	–	$-2.3(1.1)\times10^{-14}$
Ratio of Shapiro harmonics, ς	0.778(2)	
Companion mass, $m_{\rm c}$ ($M_\odot$)	–	1.299(4)
Total system mass, $M_{\rm tot}$ ($M_\odot$)	–	2.60551(19)
Derived quantities		
Rotation period, P, (ms)	41.009823581203(11)	41.009823581195(11)
First derivative of rotation period, $\dot{P}$	$4.945(2)\times10^{-20}$	$4.944(2)\times10^{-20}$
Intrinsic spin-down rate, $\dot{P}_{\rm int}$	$4.36(9)\times10^{-20}$	$4.35(9)\times10^{-20}$
Galactic longitude, ℓ	$53\overset{\circ}{.}3426(11)$	
Galactic latitude, b	$15\overset{\circ}{.}6119(12)$	
NE2001 DM-derived distance (kpc)	1.20(36)	
YMW16 DM-derived distance (kpc)	0.91(18)	
Height above Galactic plane, z (kpc)	0.24(5)	
Total proper motion, $\mu_{\rm tot}$ ($\rm mas\,yr^{-1}$)	9.52(7)	9.56(7)
Transverse velocity, $v_{\rm trans}$ ($\rm km\,s^{-1}$)	43^{+51}_{-34}	
Total peculiar velocity, $v_{\rm tot}$ ($\rm km\,s^{-1}$)	49^{+77}_{-30}	
Characteristic age, $\tau_{\rm c}$ (Gyr)	13	
Surface magnetic field strength, $B_{\rm s}$ (10^9 G)	1.44	
Mass function, f ($M_\odot$)	0.2942356(3)	0.2942355(5)
Einstein delay, γ (s)	–	0.001441
Inclination of orbit, i (deg)	–	75.8(7)*
Orthometric amplitude of Shapiro delay, h_3 (μs)	–	3.02
Pulsar mass, $m_{\rm p}$ ($M_\odot$)	–	1.306(4)†

PK parameters $\dot{\omega}$, $\dot{P}_{\rm b}$ and ς were measured using the orthometric parameterised Shapiro delay Damour-Deruelle timing model (DDH, Freire and Wex, 2010) in TEMPO2 whereas the quoted masses, γ and the inclination angle were measured and derived assuming GR as the correct theory of gravity (DDGR, Damour and Deruelle, 1986). Figures in parentheses represent the nominal 1σ (68%) uncertainties in the least-significant digits quoted. Time offsets between telescopes and different instruments were also fit for using Arecibo's PSPM backend at 430 MHz as a reference, however they are not astrophysical, so they are not shown here. Using any other backend as the basis for our jumps gave consistent results.

*Calculated using the binary mass function and the component masses in the relation $f = (m_{\rm c}\sin i)^3/M_{\rm tot}^2$. The reported uncertainty is a result of error propagation on the masses and mass function.

†Derived from $M_{\rm tot} - m_{\rm c}$.

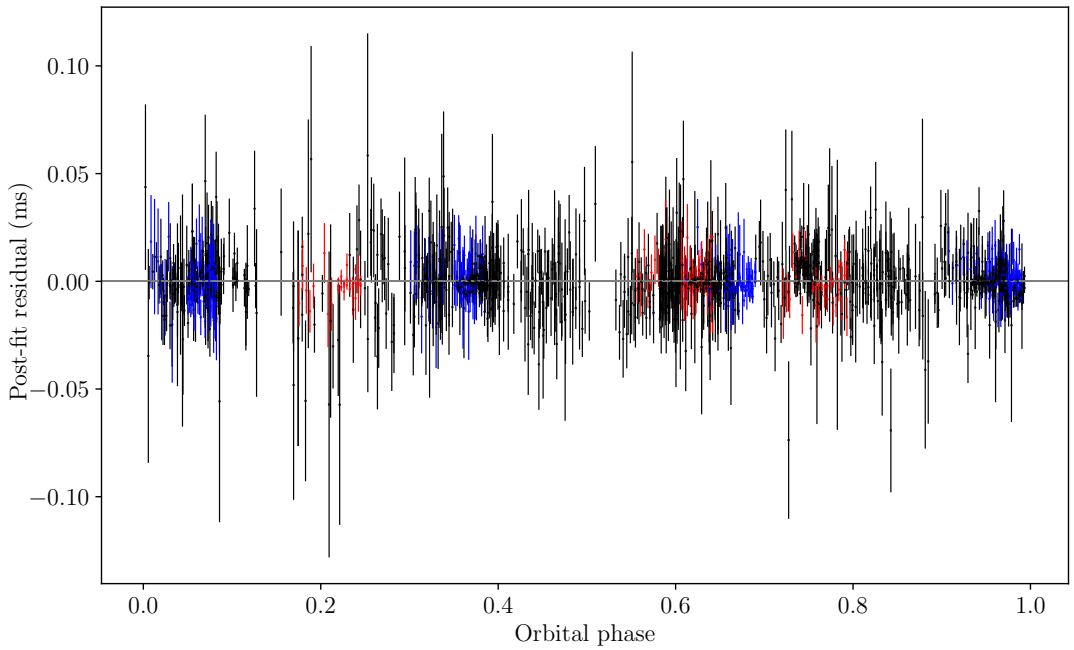


Figure 3.5: The DDGR solution post–fit timing residuals for PSR J1829+2456 as a function of orbital phase. The blue points correspond to TOAs from the first dense campaign (58402 – 58408) and the red points correspond to TOAs from the second dense campaign (58748 – 58752).

and Weisberg, 1989) has the PK parameters re-cast due to the constraints as imposed by the theory of general relativity (Equations 1.28 - 1.32). The PK parameters found by DD are not fittable in the DDGR model as $m_{\rm tot}$ and $m_{\rm c}$ are fit instead, these PK parameters can be used to compare the DD model with the DDGR model.

More recently, the DD model has been reparametrised to better probe effects of Shapiro delay on timing residuals. This DDH model (DDH, Freire and Wex, 2010) fits a new observable, ς, which is linearly related to s in DD (so as to maintain theory–independence), in tandem with higher harmonics of that observable, h_3 and h_4 (see Equations 3.1 – 3.3). The advantage of the DDH model is that the covariance between its Shapiro delay parameters is generally much lower than between r and s in the DD model for larger orbital inclinations and can therefore be used to more reliably determine mass constraints in lieu of any additional PK parameter measurement (such as γ or $\dot{P}_{\rm b}$).

$$\varsigma = \frac{s}{\sqrt{1-s^2}} \tag{3.1}$$

$$h_3 = r\varsigma^3 \tag{3.2}$$

$$h_4 = r\varsigma^4 \tag{3.3}$$

The full set of post–fit timing residuals as a function of date for the DDGR fit are shown in Figure 3.4. Plotting the residuals against orbital phase and highlighting the

dense campaigns (Figure 3.5) has shown their value in providing orbital coverage although there are still major gaps, especially around $0.1 - 0.2$ and $0.5 - 0.55$ phase. The TOAs made from the dense campaigns cover a wide range of the total orbit, which can potentially allow for concise measurements of some of the more difficult to measure orbital parameters, such as γ or any effects due to Shapiro delay. Further timing campaigns could possibly resolve this feature in the residuals.

Parameter covariance

Realistic independent measurements of certain parameters in the timing fit are sometimes impossible due to how those parameters manifest in the timing model. These *covariant* parameters get absorbed into other delays in the timing solution (e.g. r and s for head–on orbits; van Straten, 2013). In particular, frequency dependent parameters, such as $\dot{\nu}$ and DM, will depend on each other in the model to some degree and the effect of frequency on one parameter cannot always be fully disentangled from the effect of frequency on the other. Highly covariant parameters reduce the likelihood of making definitive conclusions as to whether the significance of those parameters is real or whether it is produced by other effects. From Table 3.3, calculated using the "matrix" output plugin in TEMPO2, there appears to be a high covariance between T_0 and the orbital parameters P_b and ω. This is expected due to how these parameters are updated in the model. Once a significant $\dot{\omega}$ measurement is made in a theory independent way, subsequent ω measurements in GR can be calculated from the reference epoch and the argument of periastron at that epoch, ω_0, as:

$$\omega = \omega_0 + \frac{P_b \dot{\omega}}{\pi} \tan^{-1} \left[\left(\frac{1+e}{1-e} \right)^2 \tan \frac{u}{2} \right] \tag{3.4}$$

where u is given by numerically solving Equation 3.5 at the pulsar proper time, t:

$$\frac{2\pi}{P_b}(t - T_0) = u - e \sin u \tag{3.5}$$

from Damour and Deruelle (1986), so ω can never be truly disentangled from P_b or T_0, and to a lesser extent, e. There is also expected covariance between the position and proper motion which is observed here. Extra red noise uncertainty for all of these parameters was applied using the method from Coles et al. (2011) before reporting values. The mass measurements, which are also often covariant with orbital parameters, displayed little covariance so the mass constraints can still be relied upon without need for much extra error. A good indicator for high covariance is the 'number of nines' approach where the number of consecutive nines from the decimal point in the square regression are counted. A higher 'N.o.N' value than 1 (cov ~ 0.99) implies severe covariance and is often implicit of an incorrect timing solution if too many parameters display this degree of covariance. In the case of PSR J1829+2456, no parameter obtained a 'N.o.N' score greater than 1.

Parameters that are naturally covariant, such as all parameters that share frequency dependency (e.g. $\dot{\text{DM}}$ and $\dot{\nu}$) show minor, although expected, levels of covariance in the

Table 3.3: Covariance matrix of PSR J1829+2456 for the DDGR binary fit.

Param	α	δ	ν	$\dot{\nu}$	DM	$\dot{\mathrm{DM}}$	$\ddot{\mathrm{DM}}$	$\dddot{\mathrm{DM}}$	μ_α	μ_δ	P_b	T_0	x	ω	e	$\dot{P}_\mathrm{b}^\mathrm{X}$	$\dot{x}$	m_c	M_T
α	1.000	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
δ	0.012	1.000	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
ν	0.075	0.342	1.000	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
$\dot{\nu}$	-0.089	-0.340	-0.889	1.000	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
DM	0.023	0.223	0.604	-0.602	1.000	–	–	–	–	–	–	–	–	–	–	–	–	–	–
$\dot{\mathrm{DM}}$	-0.149	-0.284	-0.957	0.958	-0.567	1.000	–	–	–	–	–	–	–	–	–	–	–	–	–
$\ddot{\mathrm{DM}}$	0.264	0.099	0.475	-0.480	0.160	-0.659	1.000	–	–	–	–	–	–	–	–	–	–	–	–
$\dddot{\mathrm{DM}}$	-0.260	-0.015	-0.213	0.218	0.023	0.419	-0.958	1.000	–	–	–	–	–	–	–	–	–	–	–
μ_α	-0.789	-0.008	-0.069	0.083	-0.025	0.142	-0.255	0.252	1.000	–	–	–	–	–	–	–	–	–	–
μ_δ	0.005	-0.988	-0.283	0.283	-0.180	0.222	-0.059	-0.012	-0.012	1.000	–	–	–	–	–	–	–	–	–
P_b	0.261	0.143	-0.171	0.147	-0.116	0.135	-0.071	0.035	-0.270	-0.132	1.000	–	–	–	–	–	–	–	–
T_0	-0.261	-0.133	0.109	-0.082	0.083	-0.075	0.031	-0.009	0.267	0.125	-0.961	1.000	–	–	–	–	–	–	–
x	0.085	0.040	-0.260	0.259	-0.124	0.257	-0.191	0.135	-0.091	-0.060	0.135	-0.076	1.000	–	–	–	–	–	–
ω	-0.252	-0.110	0.164	-0.136	0.107	-0.132	0.084	-0.052	0.259	0.107	-0.848	0.987	-0.197	1.000	–	–	–	–	–
e	-0.032	0.037	0.233	-0.239	0.123	-0.236	0.172	-0.117	0.058	-0.028	-0.129	0.100	-0.584	0.195	1.000	–	–	–	–
$\dot{P}_\mathrm{b}^\mathrm{X}$	0.029	-0.101	0.195	-0.198	0.085	-0.189	0.108	-0.063	-0.023	0.115	-0.137	-0.104	-0.335	-0.089	0.100	1.000	–	–	–
$\dot{x}$	-0.168	-0.077	0.002	-0.009	0.005	0.003	-0.002	0.002	0.155	0.068	0.096	-0.111	-0.689	-0.070	0.140	0.055	1.000	–	–
m_c	0.005	0.003	-0.337	0.329	-0.156	0.335	-0.246	0.172	-0.021	-0.035	0.255	-0.202	0.878	-0.337	-0.686	-0.375	-0.264	1.000	–
M_T	0.268	0.104	-0.141	0.116	-0.103	0.107	-0.067	0.040	-0.276	-0.092	0.667	-0.786	0.060	-0.776	-0.109	0.110	0.113	0.174	1.000

model. Further timing campaigns at different frequency bands could certainly reduce this covariance.

3.4 Results

Newly measured properties of PSR J1829+2456 have shown the system to be similar to several other DNS systems for which component masses and proper motion have been measured. Given the evolutionary relationship between the masses and the parameters discussed, it is believed that these systems underwent similar evolutionary processes (Tauris et al., 2017). Table 3.4 compares the known recycled DNS systems for which mass measurements have been made or bounded.

3.4.1 Post–Keplerian parameters

The Keplerian description of orbital mechanics is irreconcilable with observations of orbits containing sufficiently massive orbits, such as NS–WD and DNS binaries. As mentioned in Section 1.2.4, there are several “post–Keplerian" (PK) parameters all of which can be directly observed given enough orbital coverage for a system. It was the goal (and partial success) of the dense campaign observations to help constrain the Keplerian and PK parameters of the system. As such, the best timing solution has found a significant Shapiro delay measurement with $\varsigma = 0.778 \pm 0.004$ and h_3 fixed at its GR-derived value of $3.02\,\mu$s. The new timing analysis also provided better constraints for $\dot{\omega}$ leading to a more precise total mass measurement. A significant value for $\dot{P}_\mathrm{b}$ has also been measured to be $-2.9(2.4) \times 10^{-14}$, however the uncertainty in this measurement means $\dot{P}_\mathrm{b}$ cannot constrain mass measurements as precisely as $\dot{\omega}$ and ς and the measurement is, at best, a weak test for GR.

3.4.2 Mass measurements

Measuring any two post–Keplerian parameters and formulating them in terms of GR allows for constraints to be put on the component masses. In the DDGR model, the fittable mass parameters are the companion's mass and the total mass with any previously measured PK parameters held fixed. If the only measured PK parameter is $\dot{\omega}$ and the DDGR model finds a solution for the masses, it is highly likely that Shapiro delay is present in the system and that it is responsible for the component mass constraints, otherwise the masses would be unbounded in GR. This must be the case because the current estimate for γ, as quoted by both DD and DDH, is not sufficiently accurate for PSR J1829+2456 and so cannot explain the observed uncertainties. Using the theory–independently observed $\dot{\omega}$ and ς, the total and companion mass have been measured to be $2.60551 \pm 0.00038\,M_\odot$ and $1.299 \pm 0.007\,M_\odot$, respectively. Assuming the pulsar mass is equal to the total mass minus the companion mass implies $m_\mathrm{p} = 1.306 \pm 0.007\,M_\odot$. These masses are fully consistent with the values and constraints as determined by Champion et al. (2005).

It has been shown that, for sufficiently wide orbits, the derivative of the projected semi-major axis can be highly covariant with γ, which itself is covariant with the current value

for x as well as the proper motion (Ridolfi et al., 2019, equations 25 and 43 respectively). The absolute maximum contribution to $\dot{x}$ due to proper motion has been determined to be 2.6×10^{-15} lt-s s^{-1}. However, this estimate is one order of magnitude smaller than the value for $\dot{x}$ as given by the DDGR fit, which is $4\,\sigma$ significant – a rather large significance. The contribution to $\dot{x}$ due to the Lense-Thirring effect (Krishnan et al., 2020) was also found to be insignificant when compared with the uncertainty in $\dot{x}$. Because of this unexplained value of $\dot{x}$, γ cannot be reliably measured. Given the masses determined by the DDGR model, γ is predicted by GR to have a value of 1.44 ms.

In order to ensure that the GR fit obtained was the best fit possible, the post–fit χ^2 values were probed. The main setup, using arbitrary parameters p_1 and p_2, was to fix those parameters (i.e. set the fit parameter to 0) at regular intervals and run TEMPO2 with those fixed values, measuring the χ^2 of the output fit. Splaver et al. (2002) use a procedure to calculate the Bayesian joint posterior probability density. In this method, the likelihood, Π, that the i^{th} set of values for the tuple $(p_{1,i}, p_{2,i})$ gives the lowest χ^2 (χ^2_{min}) is given by a normalised Bayesian likelihood function:

$$\Pi\left(\{X\}|(p_{1,i},p_{2,i})\right) = \exp\left(\frac{\chi^2_i - \chi^2_{\text{min}}}{2}\right) \Big/ \sum_i \exp\left(\frac{\chi^2_i - \chi^2_{\text{min}}}{2}\right) \tag{3.6}$$

where X is the data set. This gives a maximum at χ^2_{min}, as desired. Calculating the posterior probability density of p_1 and p_2 given the data set makes use of Bayes' theorem (e.g. Kendall, 1994, Section 8.9):

$$\Pi\left((p_{1,i},p_{2,i})|\{X\}\right) = \Pi\left(p_1,p_2\right) \times \Pi\left(\{X\}|(p_{1,i},p_{2,i})\right) \tag{3.7}$$

where $\Pi\left(p_1,p_2\right)$ is the prior probability for p_1 and p_2 and is the product of the reciprocal differences of their extreme values. A grid of posterior probabilities can be created from this and contour lines drawn at chosen % likelihoods, as well as the corresponding probability distribution functions. For the purposes of this work, the nominal $1-3\sigma$ limits of 68.3%, 95.4% and 99.83% respectively were chosen to showcase the likelihood. This method was applied to two cases:

Case 1. The parameters of interest were m_{c} and M_{tot}. The value for m_{p} was determined as the difference between the total mass and companion mass, meaning its uncertainty is on the order of the companion mass. The prior spaces were: $1.285\,M_\odot \leq m_{\text{c}} \leq 1.310\,M_\odot$ and $2.605\,M_\odot \leq M_{\text{tot}} \leq 2.606\,M_\odot$. The grid resolution was 200×200. The contour and PDF plots can be seen in Figure 3.6.

Case 2. The parameters of interest were m_{c} and m_{p} where the total mass was calculated using the precise value for $\dot{\omega}$ in the DDH solution. This puts a much tighter constraint on the component masses as the total mass is essentially fixed to the $\dot{\omega}$ line. The prior spaces were: $1.290\,M_\odot \leq m_{\text{c}} \leq 1.310\,M_\odot$ and $1.295\,M_\odot \leq m_{\text{p}} \leq 1.315\,M_\odot$ and the grid resolution was, again, 200×200. These contours and PDFs can be seen in Figure 3.7.

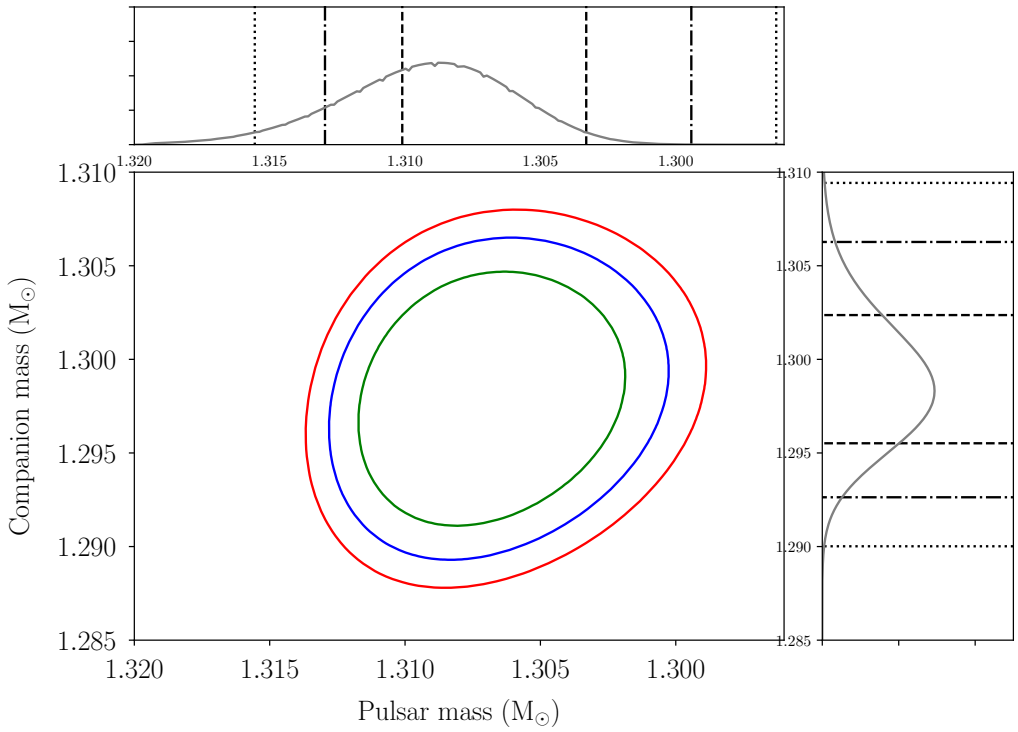


Figure 3.6: The GR–derived mass distribution for the system component masses assuming fixed M_{tot} after initial determination via GR. The vertical lines in the PDFs are the three confidence levels as determined by Gaussian variance.

It is clear that Case 2 provides a much more precise probability distribution in $(m_{\mathrm{p}}, m_{\mathrm{c}})$ space and therefore it is explored further. Since the GR–derived mass distributions as found by Case 2 are fundamentally linked to the independently observed PK parameters, as found using the DDH binary model, the contours and PK parameters can plotted to better constrain the allowed mass values. If GR is to be believed, then Figure 3.8 and its inset show the allowed mass constraints when comparing both the DDH and DDGR fit.

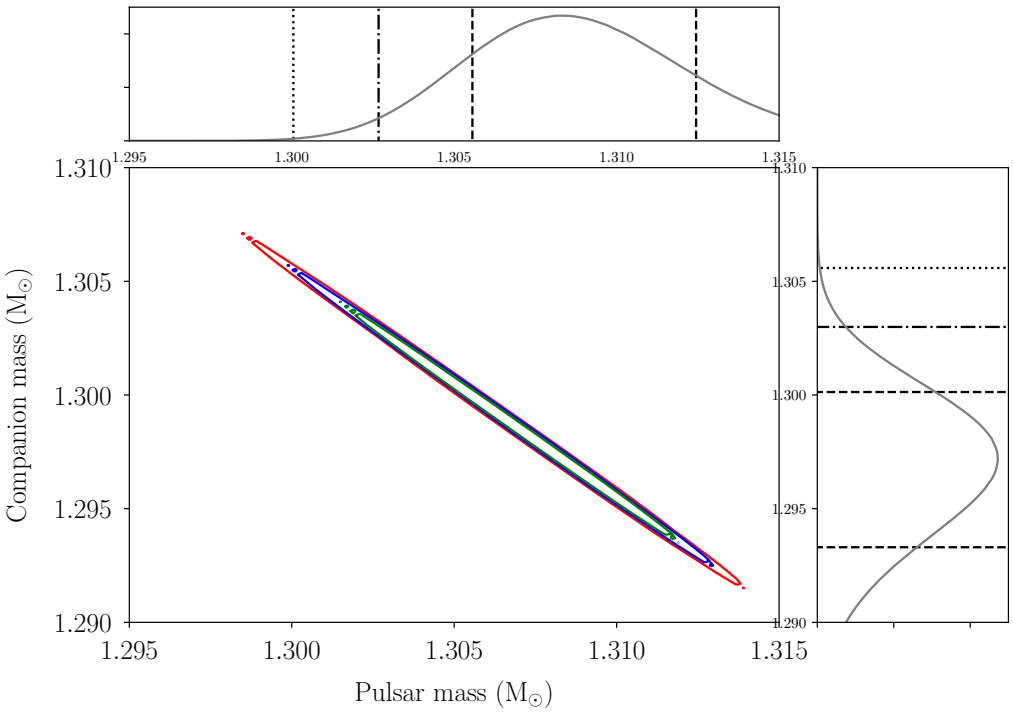


Figure 3.7: The GR–derived mass distribution for the system component masses assuming fixed $M_{\rm tot}$ as derived in a theory independent way using $\dot{\omega}$ from the Shapiro-delay harmonic Damour–Deruelle (DDH) binary model. The vertical lines in the PDFs are the three confidence levels as determined by Gaussian variance.

3.4.3 Proper motion

As a result of the latest timing campaigns, the proper motion for the system has been significantly detected in both binary models. For the DDH and DDGR models respectively, the proper motion in the right ascension, μ_α, was found to be $-5.51 \pm 0.10\,\mathrm{mas\,yr^{-1}}$ or $-5.51 \pm 0.12\,\mathrm{mas\,yr^{-1}}$ and the proper motion in the declination, μ_δ was found to be $-7.75 \pm 0.14\,\mathrm{mas\,yr^{-1}}$ and $-7.82 \pm 0.16\,\mathrm{mas\,yr^{-1}}$. This gives a total proper motion, given by the length of the additive vector of the two component proper motions, of $9.52 \pm 0.14\,\mathrm{mas\,yr^{-1}}$ or $9.56 \pm 0.14\,\mathrm{mas\,yr^{-1}}$ for the DDH and DDGR models respectively.

As mentioned in Section 3.1, calculating 3D, or even transverse, space velocity from proper motion is not often feasible for radio pulsars due to the the high uncertainties involved in DM–derived distance estimates, and very few NSs have precise parallax measurements. In order to combat this uncertainty, Monte–Carlo simulations were carried out to find both the tangential and radial components of the space velocity.

The proper motion was first converted to the co–ordinates of the "Local Standard of Rest" (LSR). This is a fictitious circular orbital path in the Galactic plane going clockwise around the Galactic centre at the current Galactocentric Solar distance, R_0. An object's velocity with respect to the LSR is known as its peculiar velocity. Even the Sun has been observed to ascend and descend through the Galactic plane and thus has a non–zero peculiar velocity which must be taken ito account (McMillan and Binney, 2010). The LSR

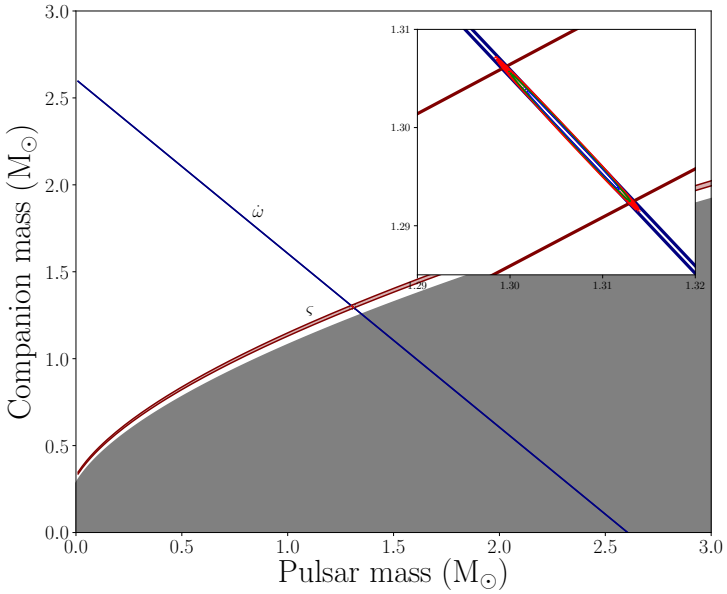


Figure 3.8: Main window: Mass-mass diagram for PSR J1829+2456 showing the GR-derived mass constraints from each PK-parameter fit. The blue region is $\dot{\omega}$ and the maroon region is ς, as reported by TEMPO2. The dashed line represents the $s = 1$ constraint. Inset: The green contoured region represents the 95% confidence region for the pulsar and and companion masses based on the DDGR model, which assumes general relativity for the timing fit.

is a good model for the general motion of object's around the Galaxy but care must be taken when interpreting LSR velocity calculations; depending on the position of the star, the co-ordinate transformation may no longer be realistic enough to disentangle from the simulated velocities' uncertainties. To convert the Earth–observed proper motion to the LSR, the method as given by McMillan (2017) was employed with solar peculiar velocity $\langle U_\odot, V_\odot, W_\odot \rangle = \langle -11.1, 12.24, 7.25 \rangle$ km s^{-1}.

For the simulations, two random priors were chosen: the DM–derived distance, d, calculated using the YMW16 electron density model, and the cosine of the orbital inclination, $\cos i$. The distance was chosen from a Gaussian distribution with the 1σ level set to be equal to 20% of the distance estimate, in accordance with the uncertainties given by Yao et al. (2017). The transverse velocity is calculated by:

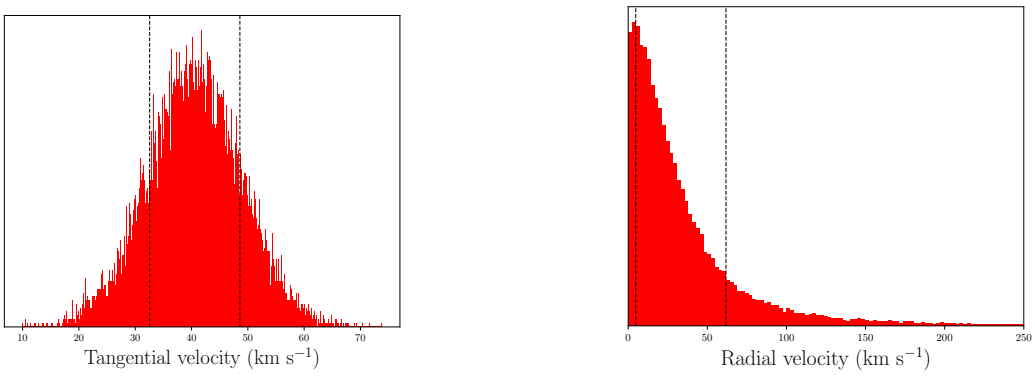
$$v_{\rm trans} = 4.74047 \mu_{\rm tot} d \tag{3.8}$$

where the scaling factor of 4.74047 is the factor that converts pc km s^{-1} to mas yr^{-1} assuming that the proper motion is small enough to use the small–angle approximation. This was done over 20,000 iterations to create a distribution of transverse velocities from which the median was taken to be the nominal value and the uncertainties were propagated from the original distance estimate and the uncertainty in the proper motion. The distribution for $v_{\rm trans}$ can be seen in Figure 3.9a. The estimated transverse velocity is then 43^{+51}_{-34} km s^{-1}.

The prior for $\cos i$ was selected from a uniform distribution between 0 and 1, giving a low–biased skew to the distribution of i. The radial velocity can be calculated by:

$$v_{\rm r} = v_{\rm trans}\cot(i) \tag{3.9}$$

where the uncertainties for $v_{\rm r}$ were linearly propagated from the uncertainty in $v_{\rm trans}$ over 100,000 iterations. 100,000 was chosen as optimal taking into account the uncertainty in the 20,000 iterations to calculate $v_{\rm trans}$. The full distribution for $v_{\rm r}$ can be seen in Figure 3.9b. Using these two velocity distributions, the total 3D space velocity was calculated as $49^{+77}_{-30}\,{\rm km\,s^{-1}}$.



(a) (b)

Figure 3.9: Distributions for the tangential (left) and radial (right) velocities for PSR J1829+2456. The dashed lines are the nominal 1σ uncertainties as obtained via error propagation from the prior uncertainties.

Until a significant parallax is observed for the system, the prospects of improving the precision on this velocity measurement are slim and the current velocity estimates are too high to draw definitive conclusions on binary evolution however the velocity estimates for PSR J1829+2456 are consistent with a low-kick evolution model.

3.4.4 Kinematic effects

The component kinematic contributions for $\dot{P}$ were then calculated. These are given by the second derivative of the line-of-sight distance from the pulsar to the Earth, or the first derivative of the Doppler factor. The transverse motion of a pulsar with respect to the SSB results in an ever–growing distance between the SSB and the pulsar, meaning that changes in pulse period, $\dot{P}$, as measured on Earth are generally not correct, displaying a secular increase (Shklovskii, 1970). The intrinsic spin–down rate due to this effect is given by:

$$\dot{P}_{\rm int} = \dot{P}_{\rm obs} - \frac{P\mu_{\rm tot}^2 d}{c} \tag{3.10}$$

where c is the vacuum speed of light. Using calculations in (Stovall and et al., 2019), three main contributions were obtained: 8.28×10^{-21} for the Shklovskii effect (Shklovskii, 1970), -1.23×10^{-21} for the difference in rotational accelerations between the Solar System and the pulsar, projected along the direction between the two, and -1.09×10^{-21} for the

difference in vertical accelerations between the Solar System and the pulsar, projected along this same direction. The total correction to the spin period is then 5.97×10^{-21}. Subtracting this from $\dot{P}$, the intrinsic $\dot{P}$ was found to be $(4.35 \pm 0.09) \times 10^{-20}$, and values for the pulsar characteristics as described in Table 3.2.

The equivalent kinematic contributions to $\dot{P}_\mathrm{b}$ are 2.0×10^{-14} (Shklovskii), -3.0×10^{-15} (rotation acceleration) and -2.6×10^{-15} (vertical acceleration difference). The total predicted kinematic contribution to $\dot{P}_\mathrm{b}$ is then $(1.5 \pm 0.2) \times 10^{-14}$. The DDH measurement for $\dot{P}_\mathrm{b}$ gives $(-2.9 \pm 1.1) \times 10^{-14}$ $(1\,\sigma)$. With the DDGR model, the contribution to $\dot{P}_\mathrm{b}$ that is in excess of the GR mass prediction, $\dot{P}_\mathrm{b}^\mathrm{X}$ (XPBDOT in TEMPO2) gave a significant measurement of $(-2.3 \pm 1.1) \times 10^{-14}$ $(1\,\sigma)$. The difference between the prediction and observation is therefore $(-3.8 \pm 1.1) \times 10^{-14}$, which is more than $3\,\sigma$ significant. As for the anomalous value of $\dot{x}$, this could be caused by systematic errors in the data, but if the effect is real, there may be some nearby mass accelerating the system. Continued timing will be necessary in order to verify this.

3.4.5 Dispersion measure

Owing to the broader frequency range of data, the dispersion measure (DM) has seen an observed improvement in precision at $13.707 \pm 0.004\,\mathrm{cm}^{-3}\,\mathrm{pc}$. Furthermore, the first, second and third order rates of change in the DM have been found to be $-0.0022 \pm 0.0008\,\mathrm{cm}^{-3}\,\mathrm{pc}\,\mathrm{yr}^{-1}$, $-0.00024 \pm 0.00008\,\mathrm{cm}^{-3}\,\mathrm{pc}\,\mathrm{yr}^{-2}$ and $(-7.1 \pm 3.2) \times 10^{-6}\,\mathrm{cm}^{-3}\,\mathrm{pc}\,\mathrm{yr}^{-3}$ respectively. The most recent timing campaigns were conducted simultaneously at two frequencies (430 MHz and 1400 MHz) for the entire length of the campaign, which in theory can allow for time–sensitive DM variations to be detected. Added to this are the subband divisions present in the L-band TOAs. DM variations from the solar system have all but been absorbed into the current emphemeris, but interstellar medium (ISM) DM variations are much harder to characterize. Pulsars in binary systems with high inclination angles can show DM variations from the envelope of the companion in a similar physical principle to Shapiro delay (Freire et al., 2003). Some of these pulsars show DM variations that are independent of a constant first derivative measurement (You et al., 2007) implying higher order corrections are necessary, so now that PSR J1829+2456 has an observed DM rate of change, this is worth investigating.

In order to probe measured variations in DM, the “Stridefit2" plugin for TEMPO2 was used. Stridefitting is the process of fitting a particular parameter (in this case the DM), whilst keeping all other parameters constant, within a small segment of time. The time segment has a user–defined ‘window size’, t, and an offset value, δt, which determines the time step to the next segment. Segments should overlap $(t > \delta t)$ to ensure minimal interpolation. One must choose a suitable t and δt to not make the resolution too fine so as to detect random noise glitches but not too coarse as to average out meaningful DM measurements. The values chosen were $t = 100$ days and $\delta t = 50$ days. Thee values were chosen based on the mode time between observing days being about 30 days, giving three observations per segment. Not all TOAs could be used in the stridefit as only dates with two frequency measurements could be used; therefore only TOAs spanning dates

57950 − 58900 were chosen for the fit with dates 58050 − 58200 omitted owing to the lack of L-band data for those dates. As one can see from Figure 3.10, the variations are rather complex with a gentle increase, consistent with the measured DM first derivative. This implies that there are no significant inexplicable trends in DM variation over this time period when accounting for the measured DM derivatives. The best fit was found using all available and valid TOAs. This method did not affect the fit from the control fit in a significant way, implying that all observed segmented DM variations can indeed be explained by the apparent DM–derivative observation.

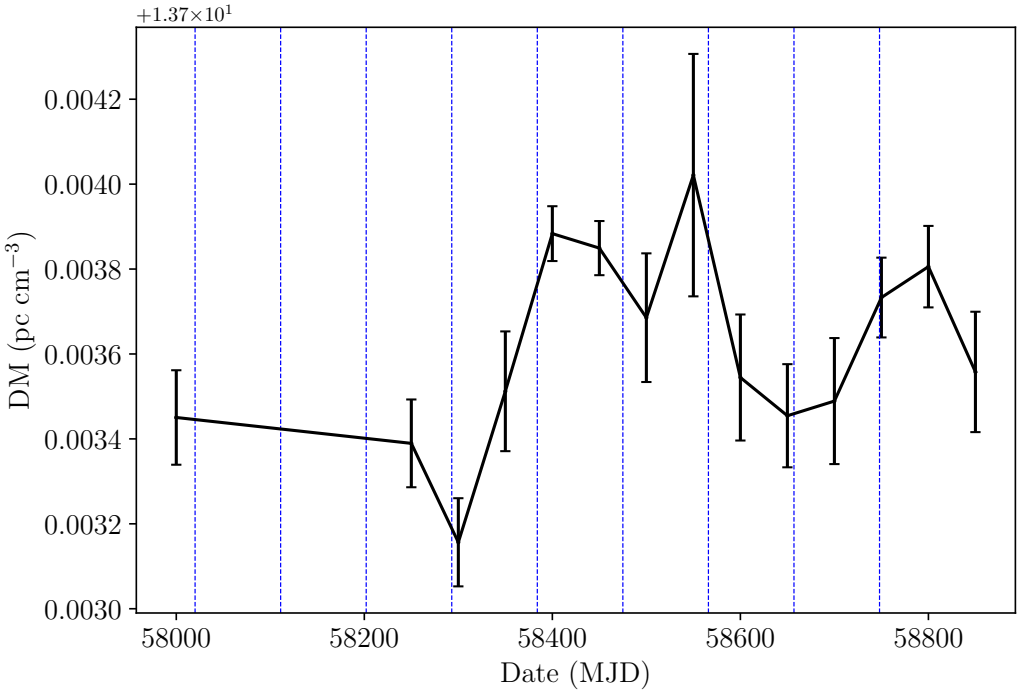


Figure 3.10: The "Stridefit2" solution for DM using TEMPO2, with dates spanning 57950 − 58900. The window around each point was 100 days with time steps of 50 days. The vertical dashed lines (blue) represent constant offsets of 3 months.

The significant measurement of higher-order DM derivatives were initially included in the model as they brought the χ^2 in TEMPO2 closer to unity. Adding more timing data did not change the nominal value of these parameters but removing them worsened the fit for the complete data set as well as subsets, however whether or not these measurements describe real phenomena remains to be seen. As previously mentioned with regards to the significant measurement of $\dot{x}$, a nearby masses could be accelerating the system which may cause non-linear propagation effects, or a small dust cloud may be periodically entering the line of sight due to the pulsar's orbit. Further work should include a DM-stridefit which is dependent on orbital phase to confirm / reject this hypothesis. These measurements may also be covariant with an unmodelled parameter (such as γ). This is likely given the results of Table 3.3 although no other parameters (with the DM derivative fits removed) gave as good a fit to the timing residuals. It should be noted however that removing the DM derivatives from the model does not significantly affect the orbital parameters or masses,

which were the main focus of this chapter.

3.4.6 Eccentricity

A more precise value for the orbital eccentricity has also been measured and compared against the population. Much like the space velocities, there is a clear increase in eccentricity with respect to the companion mass (Figure 3.15), with many currently known DNS binaries having low eccentricities. Low eccentricity implies a low mass-loss event and many of the lowly eccentric pulsars with above median companion masses have, or are predicted to have, smaller kick-velocities. This suggests a symmetric, low mass-loss SN, since the companion is similar in mass to the pulsar. These findings also agree with the theory that a larger resulting DNS eccentricity corresponds to those systems which also have undergone a large natal kick post-supernova as described in Figure 3.11.

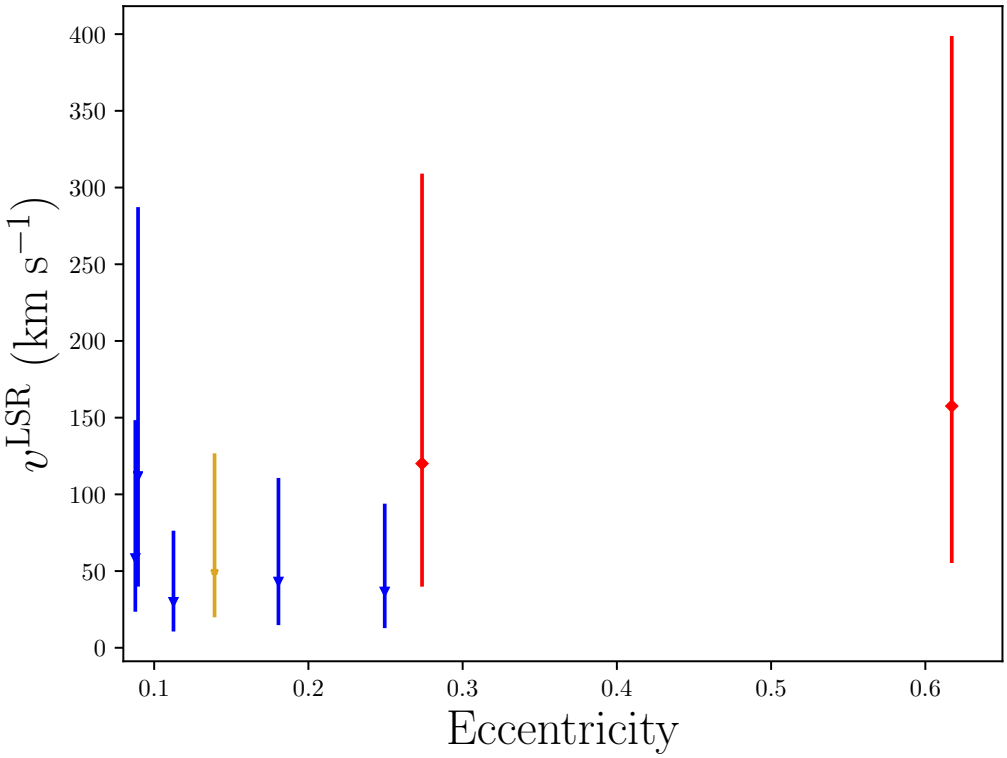


Figure 3.11: The space velocities of DNS systems with respect to the orbital eccentricity. The red (diamond) points represent those systems believed to have been formed from a violent asymmetric SN, whereas points in blue (triangles) represent the symmetric pathway. The golden (triangle) point is PSR J1829+2456.

3.5 Profile evolution

Most pulsars exhibit an increase of pulse width and profile component separation at lower frequencies, a phenomenom which has been attributed to higher frequency emission being produced closer to the NS surface (Komesaroff, 1970). This radius–frequency mapping

(Cordes, 1978) is not observed as readily for millisecond pulsars as it is for slower pulsars due to the proposed smaller size of the emission region in millisecond pulsars. In the emission picture as outlined by Komesaroff (1970), plasma flows from the NS surface along the open magnetic field lines (see Figure 1.1) and the emitted photons travel tangentially to these fields lines. The field line opening angles are dependent on the height of emission in the classic geometrical way: the closer the emission is to the surface, the smaller the opening angle of the field lines will be for the photons to emit tengentially off from. Assuming this is correct, the emission height separation can be given by the width differences between the higher and lower frequencies, $W_{\rm h}$ and $W_{\rm l}$:

$$\Delta r_{\rm em} = 2\pi c \frac{(W_{\rm h} - W_{\rm l})^2 \, (\phi_0 - \phi_{\rm l})^2 \sin^2 \alpha}{2.25 P W_{\rm l}^2} \tag{3.11}$$

from Nowakowski (2000), where α is the misalignment angle between the rotation and magnetic axes. This is by no means the only effect on observed pulse width and it has been determined that pulse width is also a function of the “slice" of the beam the observer is looking at, however assuming this does not change much over the course of one observation, Equation 3.11 holds for a single observation.

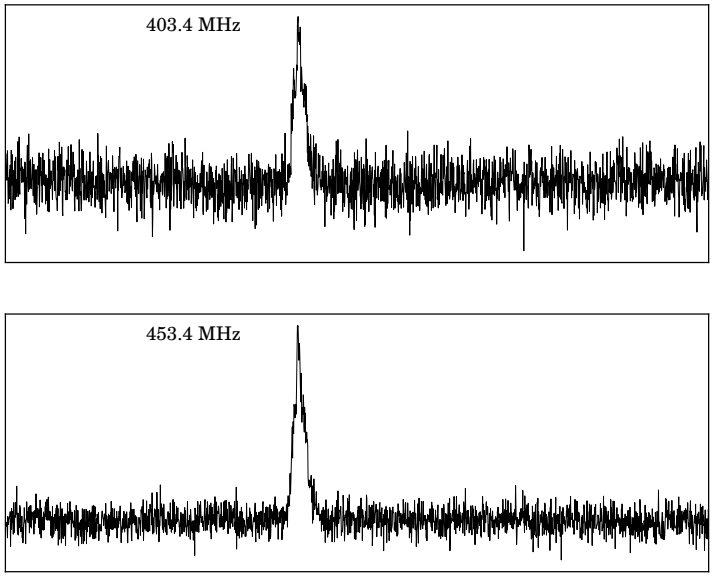


Figure 3.12: The sub-banded profiles for PSR J1829+2456 for data taken with the 430 MHz receiver at Arecibo. The date of observation was 58751 representing a total integration time of 69 minutes. These plots were made after RFI excision using most effective method and setup as determined in Section 2.2.4.

PSR J1829+2456 has been observed as having a wider profile in the 430 MHz region when compared with the L-band region, so it is worth investigating to see if subtle profile variations occur within sub–bands of frequency. Both the 430 MHz and the L-band data for MJD 58751 were divided into subbands representing 50 MHz and 200 MHz blocks

repectively, amounting to 2 frequency channels for the former and 4 channels for the latter. The sub-banding was performed after local dedispersion and strict RFI excision (Section 2.2.4). Figures 3.12 and 3.13 show the average profile shape in each frequency block.

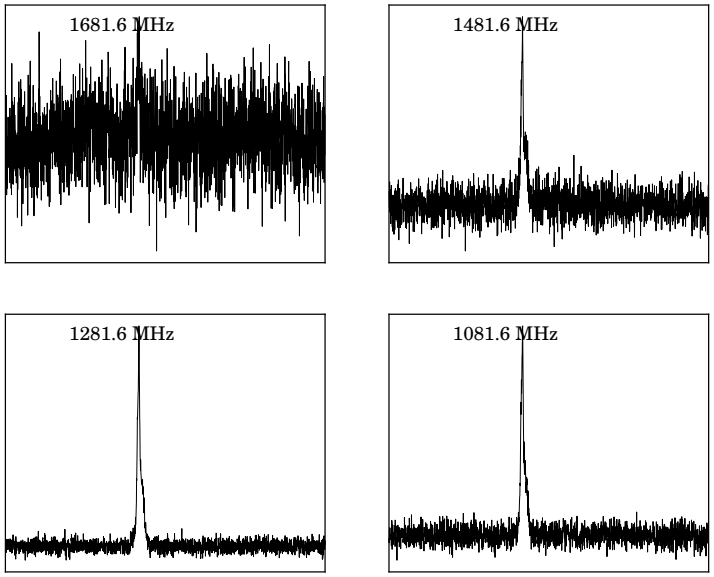


Figure 3.13: The sub-banded profiles for PSR J1829+2456 for data taken with the L-band receiver at Arecibo. The date of observation was 58751 representing a total integration time of 60 minutes. These plots were made after RFI excision using most effective method and setup as determined in Section 2.2.4.

Although the interference at high frequencies in the L–band is too great to get any discernible profile, there may be some subtle profile shape changes over both frequency bands. The phase error of each sub–banded standard profile with respect to the band-centred template was always less than one phase bin, meaning that precision timing is not affected by the profile shape change. The profiles do not get much wider or broader at lower frequencies in both bands, and a subtle change in FWHM is in line with expected frequency-dependent trends in MSPs (e.g. Pennucci, 2019). The ISM is also responsible for some profile evolution in the frequency domain (Craft and Comella, 1968) and not taking into account these profile shape changes has resulted in sub–optimal TOAs in the past (cf. NANOGrav Collaboration et al., 2015; Lentati et al., 2017), although here there is no such profile variation. Time dependent profile variation PSR J1829+2456 was not explored in detail but estimates on pulse width variations over time suggests that the profile shape has been stable across the most recent three year campaign to within the mean TOA uncertainty for that time period. This is expected as the main contributor to apparent changes in pulse shape is gyroscopic precession about the rotation axis. For PSR J1829+2456, the gyroscopic precession period is calculated to be approximately 42 centuries, so undetectable through observation at this precision. Further study on profile evolution must be conducted in order to confirm these qualitive analyses.

3.6 Evolution of the system

This system is a great resource for studies concerning DNS formation and evolution scenarios. A complete analysis of our understanding of DNS evolution is given by Tauris et al. (2017) and a more detailed introduction is given in Section 1.3, but the relevant parts to the argument of this section are reproduced, starting from the second supernova event of the helium star.

The Hulse–Taylor binary, the first DNS ever discovered, has a high 3D space velocity and a very eccentric orbit (Table 3.4). These observations fit a later theory concerning high velocity isolated pulsars as having been ejected from former binaries to suggest that supernovae imparted an asymmetric high kick velocity on the remnant NS (Bailes, 1989). This high kick velocity can be estimated from measurements of the tangential velocity which have been observed above $100\,\mathrm{km\,s^{-1}}$ in other DNSs suchs as those containing PSRs B1534+12 ($v_{\mathrm{trans}} \approx 135\,\mathrm{km\,s^{-1}}$; Fonseca et al., 2014) and J1913+1102 ($v_{\mathrm{trans}} \approx 110\,\mathrm{km\,s^{-1}}$; Ferdman et al., 2020). However many DNSs have now been observed to have low space velocities ($\lesssim 100\,\mathrm{km\,s^{-1}}$) such as those containing PSRs J1756−2251 ($v_{\mathrm{trans}} \approx 29\,\mathrm{km\,s^{-1}}$; Ferdman et al., 2014), J0453+1559 ($v_{\mathrm{trans}} \approx 25\,\mathrm{km\,s^{-1}}$; Martinez et al., 2015)[2], J1518+4904 ($v_{\mathrm{trans}} \approx 18\,\mathrm{km\,s^{-1}}$; Janssen et al., 2008), as well as the double pulsar (PSR 0737−3039A/B, $v_{\mathrm{trans}} \approx 30\,\mathrm{km\,s^{-1}}$; Ferdman et al., 2013). A revision of the theory suggests that the orbital, chemical and mass properties of NSs in a DNS lead to different evolutionary pathways (Podsiadlowski et al., 2004; van den Heuvel, 2007). Low-kick DNSs in close binaries are thought to have been formed via a mechanism of rapid electron capture onto an O-Ne-Mg or possibly iron core of the helium star in a NS-He binary. In both scenarios, the time scale for this rapid capture is much faster than the timescales for non-radial hydrodynamical instabilities to occur (Zha et al., 2019). If the core is iron, it will undergo a core collapse from all points on the surface simultaneously so the resulting supernova is believed to be symmetric. The lack of a neutrino gradient on the core surface leads to a small kick velocity in which not much mass is ejected from the system. Currently, the preferred mechanism for low-kick DNS evolution is FeCCSN (Tauris et al., 2017) as it predicts NS masses of $1.1 - 1.8\,M_{\odot}$, consistent with the current population. Furthermore, many NS masses fail to meet the constraints as determined via ECSN which predicts the DNS NS mass window to be $\sim 0.2\,M_{\odot}$ (Tauris et al., 2015). Bhattacharya and van den Heuvel (1991) show that the orbital eccentricity from a symmetric Type I supernova can be expressed as the ratio of the mass lost during the explosion to the current total system mass, although this equation assumes a roughly circular orbit:

$$e = \frac{\Delta M}{M_{\mathrm{T}}} \tag{3.12}$$

Thus systems with lower eccentricities and velocities with respect to their mass are are believed to have undergone far more stable supernovae events, only possible if the electron implosion was rapid. This is believed to be the case with J1829+2456 system. With a

[2]There is some debate as to whether this system is a DNS as opposed to a NS–WD system (Tauris and Janka, 2019)

space velocity of $\sim 50\,\mathrm{km\,s^{-1}}$, it is traveling faster than the other binaries in the "symmetric supernova camp" however it has medium-to-low eccentricity (≈ 0.139) and the mass of the companion has been found to be in a medium-to-high range with respect to these other systems (See either Figure 3.14 or 3.15). This follows known trends about companion mass and velocity (van den Heuvel, 2007) although the companion mass for PSR J1829+2456 is quite a bit higher than their prediction that second-formed NSs should have low masses of around $1.25(6)\,M_\odot$. Table 3.4 documents all of the Galactic DNSs with known orbital or proper motion parameters and Figures 3.14 and 3.15 show the population of velocities and eccentricities respectively for Galactic DNSs with known companion masses.

From Equation 3.12, the change in mass for the J1829+2456 system after the second supernova event is calculated to be $(0.3625 \pm 0.0004)\,M_\odot$ leading to a core-progenitor mass of $(1.662 \pm 0.007)\,M_\odot$, putting it at the lowest end of estimated He star core-progenitor masses (Woosley, 2019), although it cannot be understood to be the true progenitor mass owing to PSR J1829+2456's relatively large eccentricity, so it is best to consider this a ballpark figure. This mass and the systemic velocity estimates are consistent with predictions made for the system via simulation (Tauris et al., 2017, Figure 33). Specifically in the case of DNS systems, the minimum total accretion mass to the pulsar (from the companion pre-second SN) can be determined via the current spin period and can act as a measure of the extent of recycling. This is approximated by (Tauris et al., 2017):

$$\Delta M \simeq 0.22 \sqrt[3]{\frac{M_{\mathrm{NS}}}{P^4}} \tag{3.13}$$

with P measured in ms. Using this estimate, the minimum accretion mass is calculated to be $(1.700 \pm 0.007) \times 10^{-3}\,M_\odot$. This lies well within the $(0.2 - 4) \times 10^{-3}\,M_\odot$ range observed with the total DNS population.

The relatively wide orbit of this system ($P_b \approx 1.176$ days) along with the low estimated mass-loss imply that its orbital parameters have not changed much since formation (Tauris et al., 2017, Section 8.9). The companion mass of $1.299\,M_\odot$, as well as the mass ratio $q = 0.995 \pm 0.007$, support this further as they imply very little mass transfer to the pulsar during the X-ray emitting accretion stage.

Another evolutionary scenario based on neutrino driven kick (Janka, 2013) was ruled out due to PSR J1829+2456's surface magnetic field being 1.44×10^9 G – about six orders of magnitude too low to fit either neutrino kick formation scenario. This pathway of evolution is mostly saved for magnetars in binary systems.

Table 3.4: Parameters for various DNS systems in which the pulsar is the recycled NS. This list does not include systems in globular clusters, which were likely formed via exchange encounters.

PSR*	P (ms)	$P_{\rm b}$ (days)	e	Companion mass ($M_\odot$)	$\mu_{\rm tot}$ (mas yr^{-1})	d (kpc)$^\times$	$v^{\rm LSR}$ (km s^{-1})†
J0453+1559[1]	45.8	4.072	0.113	1.174(4)	7.997	0.52	29^{+44}_{-19}
J0509+3801[2]	76.5	0.380	0.586	1.46(8)	–	7.08	–
J0737−3039A[3]	22.7	0.102	0.088	1.2489(7)	3.885	1.17	55^{+86}_{-36}
J1411+2551[4]	62.5	2.616	0.170	> 0.92	~ 12	1.13	85^{+120}_{-51}
J1518+4904[5]	40.9	8.634	0.249	$1.05^{+1.21}_{-0.11}$	8.512	0.96	36^{+55}_{-22}
B1534+12[6]	37.9	0.421	0.274	1.3455(2)	25.34	0.93	120^{+184}_{-78}
J1753−2240[7]	95.1	13.638	0.304	–	–	6.93	–
J1756−2251[8]	28.5	0.320	0.181	1.230(7)	5.928	0.95	42^{+63}_{-25}
J1757−1854[9]	21.5	0.183	0.606	1.3946(9)	–	19.6	–
J1811−1736[10]	104.2	18.779	0.828	> 0.93	–	10.16	–
J1829+2456	41.0	1.176	0.139	1.299(7)	9.560	0.91	49^{+77}_{-30}
J1913+1102[11]	27.3	0.206	0.090	1.27(3)	9.286	7.14	112^{+175}_{-73}
B1913+16[12]	59.0	0.323	0.617	1.389(1)	1.404	5.25	157^{+242}_{-100}
J1930−1852[13]	185.5	45.060	0.399	> 1.30	–	2.48	–
J1946+2052[14]	17.0	0.078	0.064	> 1.18	–	3.51	–

*References: (1) Martinez et al. (2015), (2) Lynch et al. (2018), (3) Kramer et al. (2006), (4) Martinez et al. (2017), (5) Janssen et al. (2008), (6) Fonseca et al. (2014), (7) Keith et al. (2009), (8) Ferdman et al. (2014), (9) Cameron et al. (2018), (10) Corongiu et al. (2007), (11) Ferdman et al. (2020), (12) Weisberg and Huang (2016), (13) Swiggum et al. (2015), (14) Stovall et al. (2018).

$^\times$Distances used were derived from the YMW16 Galactic free electron distribution model (Yao et al., 2017) with DMs found using the ATNF Pulsar Catalogue (Manchester et al., 2005) except in the case of PSR J1756−2251, where the distance is given by Ferdman et al. (2014).

†Median $v^{\rm LSR}$ and 2σ (95% confidence level) errors were calculated using the Monte-Carlo method described in section 3.4.3 and rounded to the nearest integer.

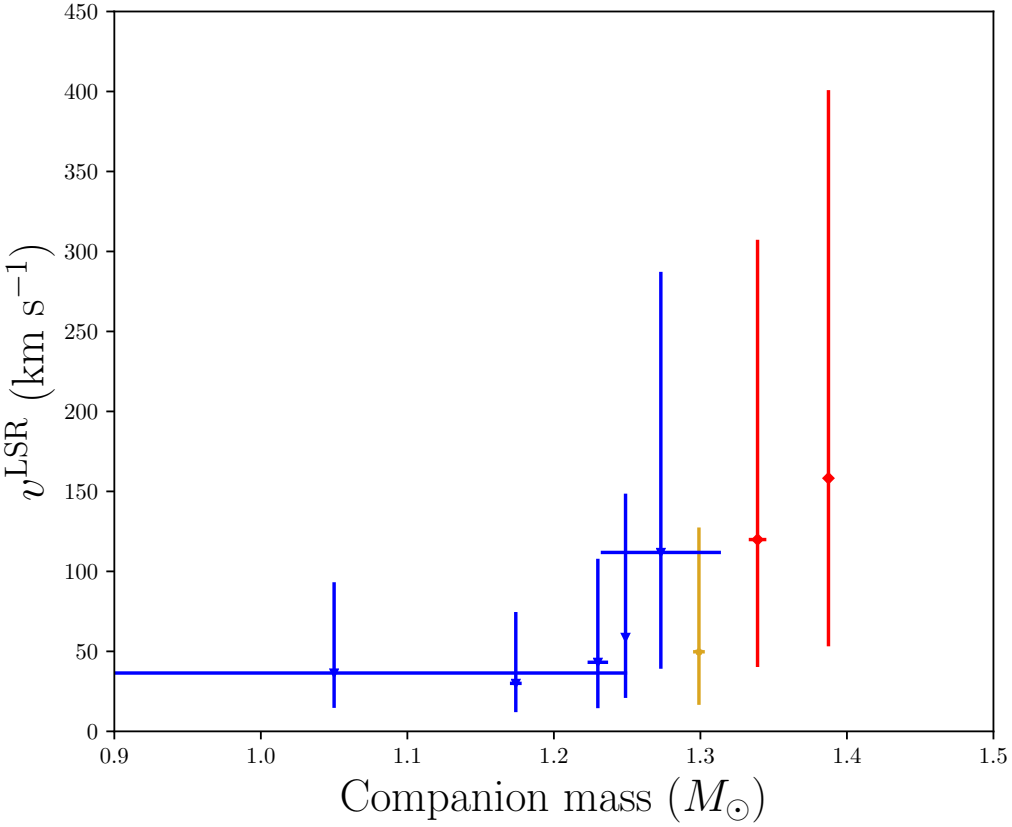


Figure 3.14: Companion mass compared with the estimated 3D velocities for Galactic DNS systems. The blue points represent those systems that are thought to have undergone a rapid electron capture onto an iron or possibly an O-Ne-Mg progenitor core leading to a low-kick, symmetric supernova. The red points represent systems theorised to have undergone a high-kick, asymmetric supernova.

3.7 Future prospects

The goal of follow up timing for PSR J1829+2456 will be to further constrain the PK parameters, the proper motion and the DM. It is estimated that, within two more years of data, including at least one carefully organised dense campaign, a significant and accurate value for γ will be measureable. Adding to this, the relatively short distance to the system with respect to other DNSs mean that measuring the parallax may soon become possible, although it is not expected to be measureable to within 0.3 mas until five more years of observation. Alternatively, a VLBI distance measurement is possible now if the YMW16 distance is close to the true value. The author suggests dual frequency follow up timing campaigns for two years to investigate these predictions. Although almost all of the TOAs obtained for this pulsar were conducted with Arecibo, that is unfortunately no longer an option[3], so other telescope options must now be explored for follow up observations.

[3]See Witze (2020) for an informative but ultimately upsetting read.

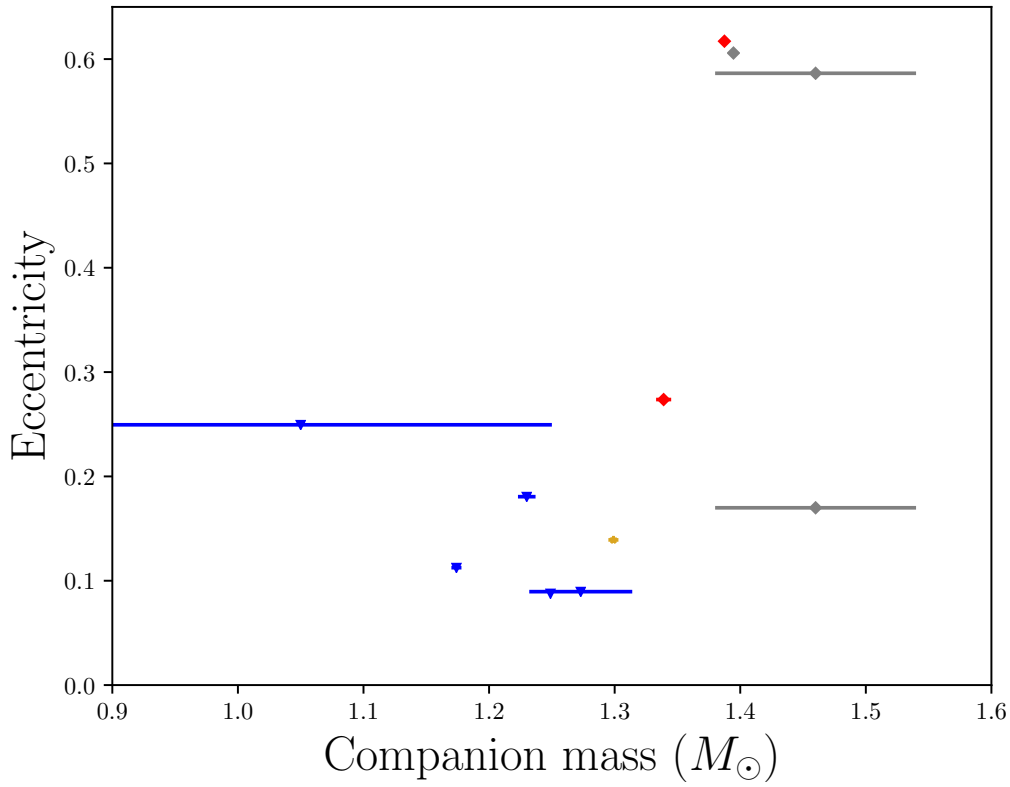


Figure 3.15: Companion mass compared with the eccentricity of Galactic DNS systems. The blue points are the symmetric supernovae systems, the red points are the asymmetric and the grey points are currently un-determined, however it is likely that they are both asymmetric supernovae remnants. Errorbars for the eccentricity are smaller than the points themselves and have been omitted for this reason.

Chapter 4

Timing of four millisecond pulsars discovered by the PALFA survey

The discovery and timing of radio pulsars is essential in order to place ever tighter constraints on NS physics, their population, formation and evolution. Due to the effect of increased dispersion and scattering from the ISM in the Galactic plane and the generally steep spectral nature of pulsars, many pulsars in the Galactic plane appear extremely faint, so a middle-range frequency ($\sim 1\,\mathrm{GHz}$) is the best option: not too large as to whiten the pulsar signal out, but not too low as to redden the signal with ISM dispersion. The known pulsar population is increasing rapidly with the help of surveys being conducted at major observatories around the world, many of which operate at $\sim 1\,\mathrm{GHz}$. Several recent and prominent pulsar surveys include: the Pulsar Arecibo L-band Feed Array (PALFA; Cordes et al., 2006), the High Time Resolution Universe survey at Parkes (HTRU; Keith et al., 2010), the Arecibo all-sky 327 MHz drift pulsar survey (AO327; Deneva et al., 2013), the Green Bank Telescope drift scan survey (Boyles et al., 2013), the Green Bank Northern Celestial Cap survey (GBNCC; Stovall et al., 2014) and the GMRT High Resolution Southern Sky survey (GHRSS; Bhattacharyya et al., 2016). These surveys have given rise to discoveries of several interesting DNSs such as: PSR J1757−1854 from HTRU (Cameron et al., 2018), which has provided three separate tests for general relativity through measurement of all five post-Keplerian parameters; PSR J1913+1102 from PALFA (Ferdman et al., 2020), whose component masses are the most asymmetric among the compact NS binary population; and PSR J1946+2052 from PALFA (Stovall et al., 2018), which has a short (4.4-hr) orbital period and is the most relativistic system discovered – it is expected to merge in only 46 Myr. As well as new DNS discoveries, many NS-WD binaries have been discovered and characterised. These are especially important in probing orbital effects on relatively circular orbits and offering insight into the formation of these binary systems.

Pulsars in binaries with other compact objects such as NSs can provide the most stringent tests of GR (Kramer et al., 2006) and other theories of gravity (Freire et al., 2012). Most relativistic binaries are expected to be found in the low-latitude Galactic plane region (e.g. Belczynski et al., 2002) where stellar populations are more numerous and generally older. It is therefore imperative that at least a handful of pulsar surveys

should be focused on this region. In this chapter, I will introduce the PALFA Galactic plane survey in more detail and the discovery of four pulsar systems. I will then discuss the follow-up timing and analysis of these pulsars and the results that arise from these observations.

4.1 The PALFA survey

The Pulsar Arecibo L-band Feed Array (PALFA) survey began in 2004 and used the seven-beam ALFA receiver at a centre frequency of 1375.5 MHz with a bandwidth of 322.6 MHz, in order to search for radio pulsars that lie close to the Galactic plane ($|b| < 5°$) (Cordes et al., 2006; Lazarus et al., 2015). It was designed to be the deepest pulsar search to date in the Galactic disk, since the thin disk is host to the majority of young stars, and therefore NSs. PALFA surveyed within two regions of Galactic longitude: the inner Galactic region ($32° < l < 77°$), with integration times of 268-s, and the outer Galactic region ($168° < l < 214°$), with 180-s integration times (Parent et al., 2019). PALFA generally searched the incoming signal for pulsars using three methods: The “Quicklook" pipeline (Stovall, 2013) processes a lower-resolution version of the data, and is often used to find particularly bright pulsars and can be applied in quasi-real time, however weaker pulsars are not typically detected via this method. The PRESTO pipeline (Ransom, 2001) is used for full-resolution searches in either Fourier space, which also uses acceleration searches to account for the Doppler effect due to pulsars in potential binary orbits, or in the time domain, by making use of so-called Fast-Folding Algorithms (e.g. Parent et al., 2018); this method is sensitive to weaker radio sources and also has a single-pulse mode for detection of RRATs and FRBs (Patel et al., 2018). The final major search method used in PALFA is Einstein@Home (Anderson et al., 2006), a distributed-processing project combining the resources of millions of home computers worldwide. Einstein@Home also processes data from the LIGO gravitational-wave detectors (Steltner et al., 2021) and the Fermi gamma-ray satellite (Pletsch et al., 2013) as well as from the PALFA survey (e.g. Allen et al., 2013).

The PALFA collaboration has discovered a total of 203 pulsars (as of April 2021), 43 of which have observed spin periods < 100 ms, potentially ideal for probing binary character due to the high chance of previous recycling from a companion star (see Section 1.3 for details). The general survey statistics (Figure 4.1) show a roughly Gaussian spread in DM with median 224 $\mathrm{pc\,cm^{-3}}$ and a logarithmically linear decay in spin period discoveries. Treating this distribution as a heavily skewed Gaussian, about 20.5% of PALFA discoveries should have spin periods $\lesssim 30$ ms, classifying them as true MSPs; indeed, 40 out of the 203 ($\sim 20\%$) do have this property. This is in contrast with the proportion of all of the discovered Galactic MSPs with $P < 30$ ms which amounts to 9% of the total pulsar population (using data from Konar and Chahal, 2019). This could be an observational selection effect due to the narrow Galactic latitude window that PALFA uses, but it may also hint at an astrophysical selection criteria governing NS populations. In any case, due to this high number of MSPs discovered, PALFA has so far discovered many pulsars in binary systems, with three having been identified as DNS systems (van Leeuwen et al., 2015; Lazarus et al., 2016; Stovall et al., 2018).

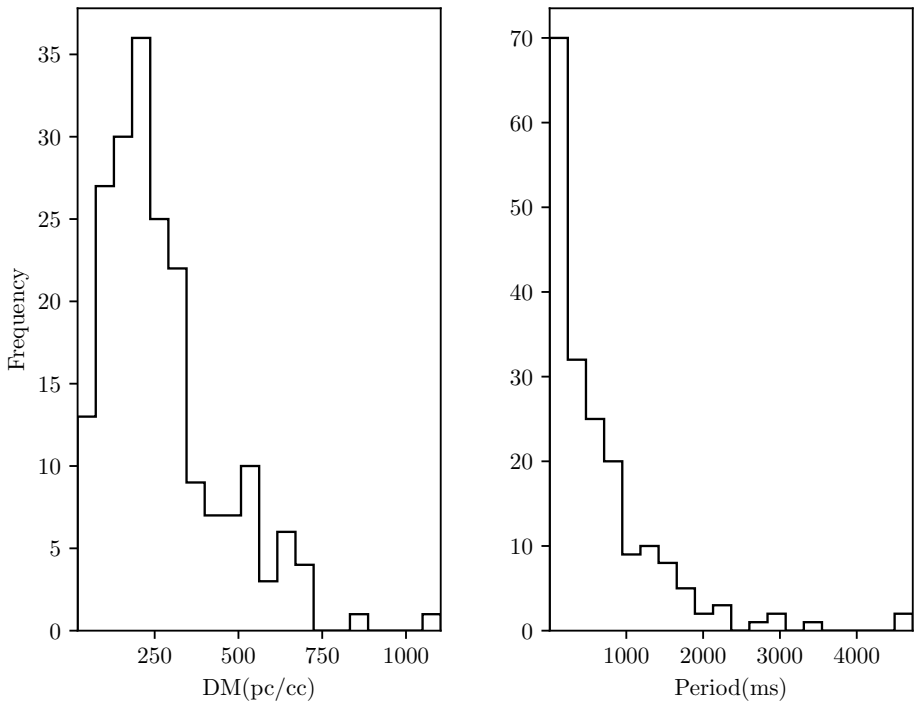


Figure 4.1: The current (as of April 2021) spin period (in ms) and dispersion measure (in $\mathrm{pc\,cm^{-3}}$) population statistics for confirmed PALFA discoveries.

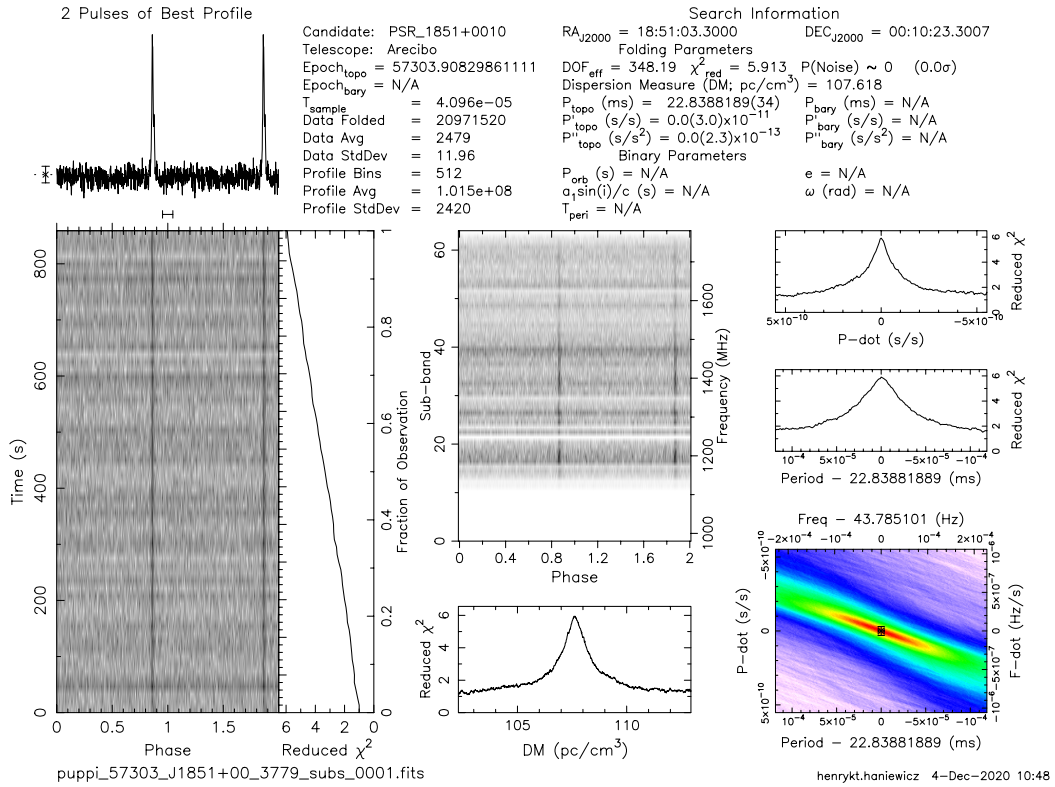
4.2 Discovery

All the PALFA pulsars analysed for this work were observed in the inner Galactic region using the 305-m Arecibo radio telescope at the full 268-s integration time using the PUPPI data acquisition backend in incoherent search-mode.

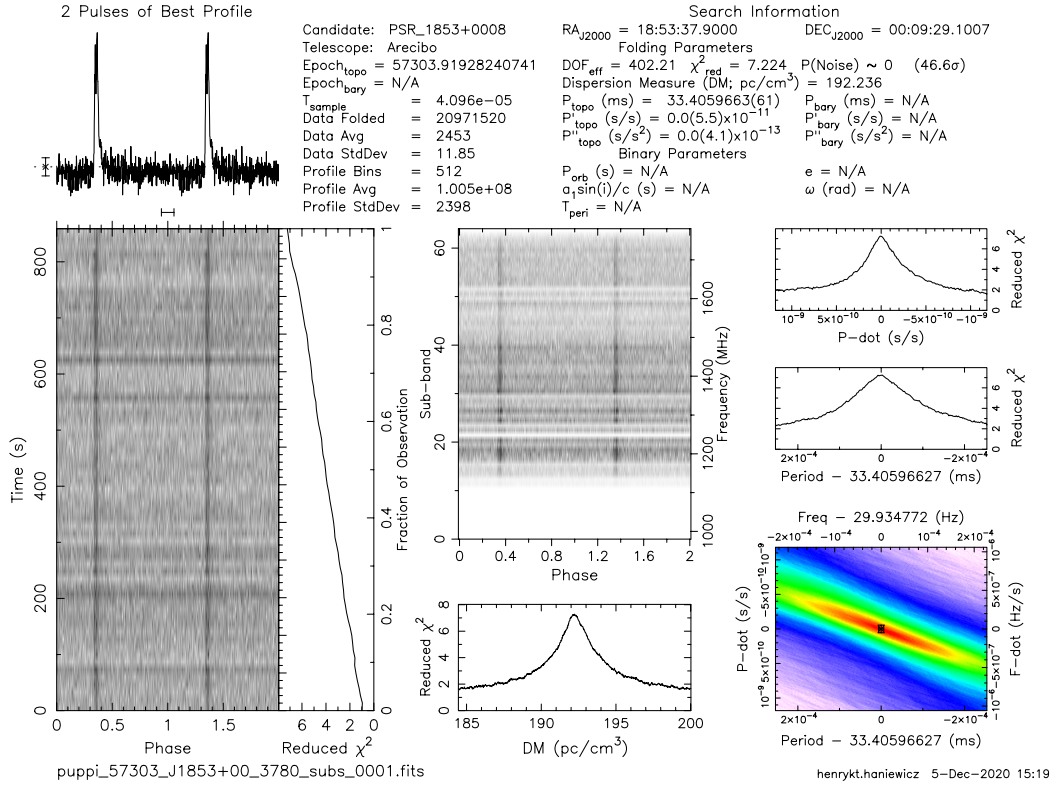
PSR J1851+0010 and PSR J1853+0008 were discovered by Einstein@Home, as part of the Binary Radio Pulsar (BRP4) search[1], in data taken on MJDs 57166 and 57168 (May 2015), respectively. PSR J1936+1805 and PSR J1936+2142 were discovered using the Quicklook pipeline on MJDs 57233 (July 2015) and 57868 (May 2017), respectively. The discovery plots for each pulsar are given in Figures 4.2 and 4.3. These are displayed as the output from `prepfold` in the PRESTO suite after refolding from the initial solution.

PSR J1851+0010 is a pulsar with a spin period of 22.8 ms and DM of $107.6\,\mathrm{pc\,cm^{-3}}$, found in the constellation Aquila. PSR J1853+0008 is a pulsar with a spin period of 34.4 ms and a DM of $192.2\,\mathrm{pc\,cm^{-3}}$, also found in the constellation Aquila. PSR J1936+1805 is a pulsar in the constellation Sagitta with a spin period of 58.3 ms and a DM of $126.1\,\mathrm{pc\,cm^{-3}}$. PSR J1936+2142 is a pulsar with a spin period of 31.59 ms and a DM of $74.8\,\mathrm{pc\,cm^{-3}}$, found in the constellation Vulpecula. Using the YMW16 Galactic electron density model (Yao et al., 2017), these DM measurements give respective distances of 3.0, 3.9, 3.6 and 3.0 kpc.

[1]For a complete catalogue of the Einstein@Home BRP4 pulsar discoveries, refer to einsteinathome.org/radiopulsar/html/BRP4_discoveries

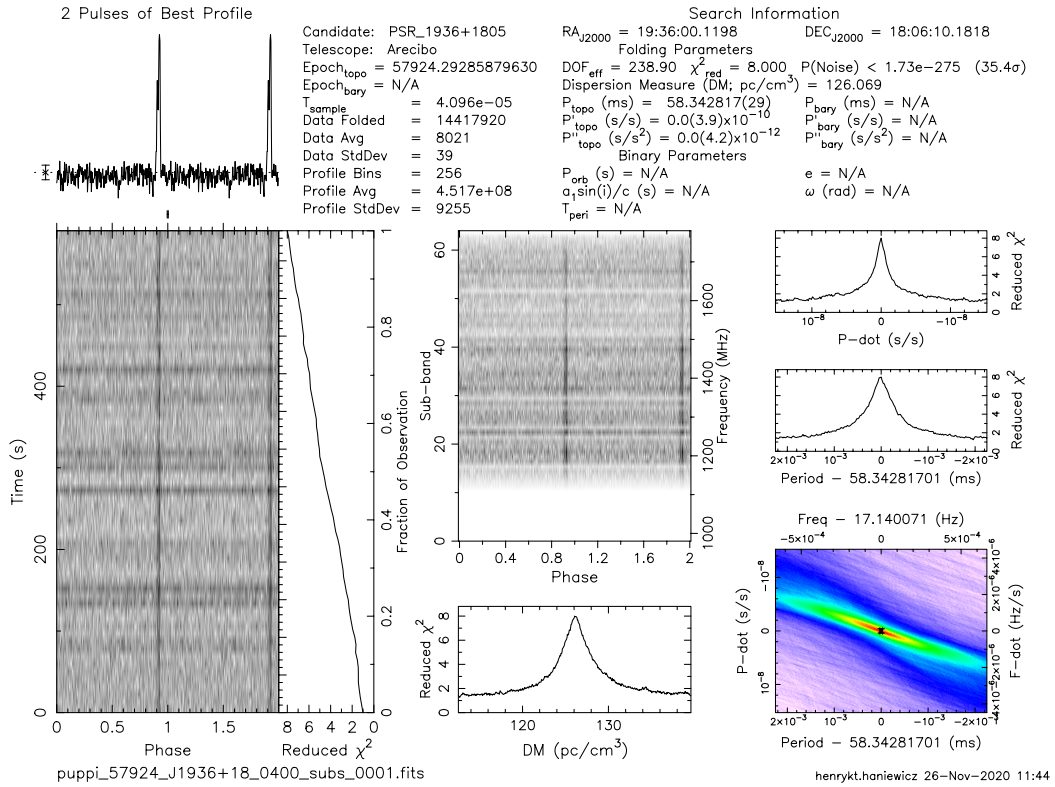


(a) PSR J1851+0010. Top: basic information of the pulsar. Right: the statistical liklihood that the folding period and its derivative are correct. Middle: the DM fold. Left: the best profile from the folded profiles (duplicated for clarity).

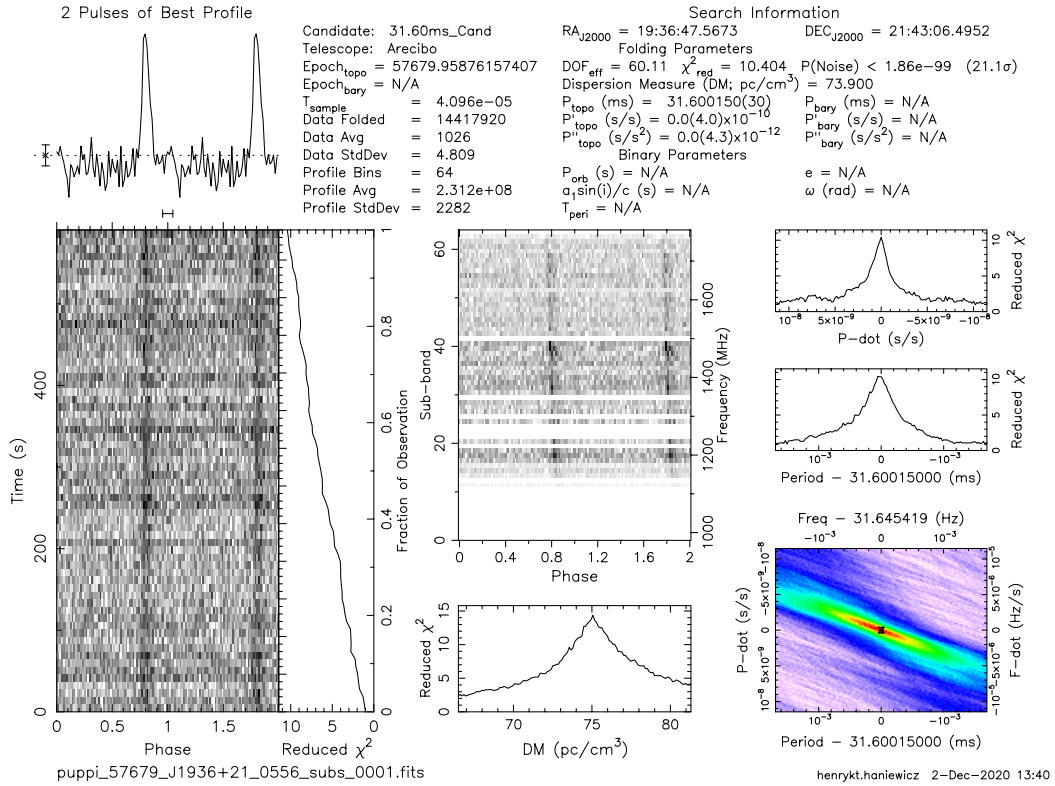


(b) PSR J1853+0008

Figure 4.2: Discovery profiles for PSRs J1851+0010 and J1853+0008, found using the Einstein@Home project.



(a) PSR J1936+1805



(b) PSR J1936+2142

Figure 4.3: Discovery profiles for PSRs J1936+1805 and J1936+2142, found using the Quicklook pipeline.

4.3 Follow-up timing

Follow-up incoherent-dedispersion observations for PSR J1851+0010 and PSR J1853+0008 ran from 57222 – 57315 and on 57341, initially once monthly for approximately 10 minutes per observation, followed by a dense campaign conducted from MJD 57311 – 57330 in 10 – 15-minute observations per epoch.

PSR J1851+0010 was also observed approximately monthly in coherent fold-mode between MJD 57330 – 58680 (November 2015 – August 2019) at a centre frequency of about 1400 MHz over 800 MHz of total bandwidth. Most observations were conducted in a ~ 15 minute window except for those observations taken on MJDs 57556 and 57558, which were each conducted over a duration of approximately 40 minutes. A linear polarisation basis was used for the receiver feeds leading to four polarisation channels in AABBCRCI format (see Section 1.1.1). Noise source calibration observations were conducted for 90 s for each pulsar at an oscillation frequency of 25 Hz and a duty cycle of 0.5 in-pulse phases (this method is identical to the noise source calibration outlined in Section 3.2). The time series data were initially folded online at Arecibo using the discovery pulse frequency of 43.785206579 Hz, a dispersion measure of 107.6727 pc cm^{-3}, divided into 512 frequency channels of width 1.5625 MHz and into subintegrations representing 10.25 s of time. To obtain more precise coherent solutions for PSR J1851+0010, the search-mode data for that pulsar were refolded at a 2048-bin resolution after a solution was found with the fold-mode data.

PSR J1853+0008 was sparsely observed in coherent fold-mode between MJD 57362 – 58680, mostly during the same session as PSR J1851+0010, at a centre frequency of 1380 MHz and over 600 – 800 MHz of bandwidth (depending on the observing day). Noise source calibration was administered as above and the data was folded online at the discovery pulse frequency of 29.93763513 Hz and a dispersion measure of 192 pc cm^{-3}. Much like PSR J1851+0010, this fold-mode data was divided into 448 or 512 frequency channels of 1.5625 MHz and into subintegrations of 10.25 s.

To find an initial solution, both pulsars were folded at a resolution of 256 profile bins; these were then refolded using the initial solution ephemeris at 512 profile bins each (representing 4.096×10^{-5} s per bin) for further timing analysis. For each of these two pulsars, one final incoherent search observation was made on MJD 59014. These were not included in the initial phase connection process due to their vast distance in time from the original search-mode data, but were folded with the initial solution's ephemeris and subsequently included in the set of refolded TOAs.

Follow-up incoherent search-mode observations were conducted from MJD 57924 – 58895 for PSR J1936+1805 and from MJD 57595 – 58680 for PSR J1936+2142. Within these observing periods, dense campaigns for PSRs J1936+1805 and J1936+2142 were conducted on MJD 58330 – 58359 and MJD 57670 – 57716, respectively. Initial incoherent folds were made for both pulsars at 128 and 64 bins per profile, respectively, using `prepfold` in PRESTO before an initial timing solution was found. After this initial solution was found, the search data for PSR J1936+1805 and PSR J1936+2142 were refolded at 512-bin resolution to obtain more precise TOAs and parameter measurements. The integrated

pulse profiles for each pulsar, following this complete method, are displayed in Figures B.3, B.5, B.8 and B.10 in Appendix B.1.

4.4 Timing analysis

Standard template profiles (Figure 4.4) for each pulsar were created for each observing mode to ensure optimal cross-correlation during TOA calculation. For the search-mode data, template profiles were created by fitting up to two Gaussian functions to the days with the highest S/N in the integrated profile. For coherent fold-mode data, template profiles were made by first administering polarisation calibration via the 90 second noise-diode pulsar calibrators and then performing flux calibration using the stably polarised QSO B1442 as a continuum calibration source (see Section 2.3). This was followed by RFI excision via a mix of CNN training with (ς, σ) architecture (see Chapter 2) and manual zapping of the raw data. The profile data were then dedispersed, aligned with the current ephemeris, and finally added together over many days. The templates are of much higher S/N than their search-mode counterparts. This means that, on average, the coherent fold-mode TOAs have smaller RMS residual error. Multiple Gaussians were then fit to the averaged profiles for all four pulsars using the Gaussian fitting routines in PSRVoid (with off-pulse regions determined by PyPulse; Lam, 2017). All resulting templates were kept at their original phase. For templates made using Gaussian fitting, the individual Gaussian components are given in Appendix B.1.

TOAs created from seperate template profiles are fundamentally different, due to the template-dependent cross-correlation process and therefore cannot be directly compared. This requires the fitting of arbitrary phase offsets between sets of TOAs that were created with between different template profiles. These 'jumps' are fit in TEMPO and TEMPO2 with respect to a reference set of TOAs. In every case throughout this chapter, the coherent fold-mode TOAs were jumped with respect to the search-mode TOAs.

The search-mode template for PSR J1851+0010 was made after RFI excision by fitting three Gaussian functions to the integrated profile for MJD 57251 at a resolution of 512 profile bins (for the refold). The fold-mode template for PSR J1851+0010 was made first through calibration and RFI excision and then by profile addition of sixteen observations spanning dates $57330 - 58680$ also at a bin resolution of 512. The search-mode template for PSR J1853+0008 was made by fitting two Gaussian functions to the integrated profile for MJD 57303 at a resolution 512 profile bins (for the refold). The fold-mode template for PSR J1853+0008 was made first through the same method of calibration and RFI excision as previously described, follow by the addition of eight observations spanning dates $57362 - 58680$. The only templates made for PSRs J1936+1805 and J1936+2142 were search-mode templates (reflecting the unitary data type) made from dates 57924 and 57679 at 256 and 64 profile bin resolution respectively. These were both fit to a two-Gaussian function.

Phase-connection for PSR J1851+0010 and PSR J1853+0008 was achieved by first phase connecting around MJDs 57556 and 57558 until as many dates as possible were phase-connected. For PSR J1936+1805, prelimary phase connection was first made for the

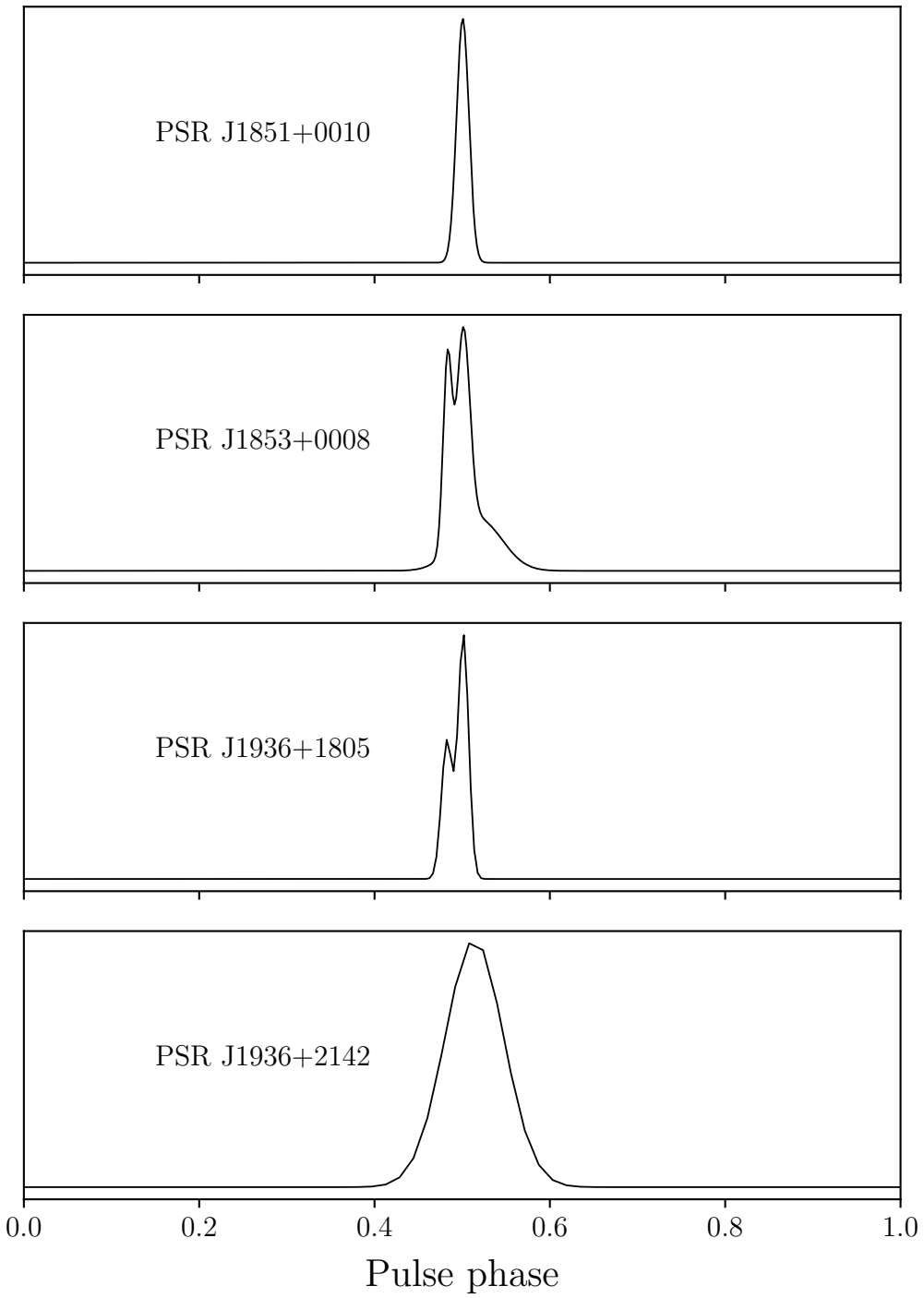


Figure 4.4: The noise-free standard profiles for PSRs J1851+0010, J1853+0008, J1936+1805 and J1936+2142. Those displayed here are the coherent fold-mode data templates for PSRs J1851+0010, J1853+0008, constructed using all available fold-mode data; and the incoherent search-mode templates for PSRs J1936+1805 and J1936+2142, constructed from the integrated profile from observation dates 57924 and 57679, respectively.

dense campaign TOAs in isolation. These now phase-connected sets of TOAs were added to their respective full TOA set and, in conjuction with the phase connected dense campaign ephemeris, were used as the starting point for further phase connection. In the case of PSR J1936+2142, a phase-connected solution was first found for the dense campaign. The partial solutions were used as an initial starting point for DRACULA (Freire and Ridolfi, 2018) to find a phase-connected solution for each full data set.

DRACULA attempts to phase-connect TOAs by first starting with completely unconnected TOAs separated by jumps which are designed to ensure coherence. Starting from densely populated regions (i.e. dense campaigns), DRACULA removes one arbitrary jump at a time and updates the timing model to reflect the jump removal by attempted to reduce the reduced-χ^2 to unity. One-by-one, more observations are linked together in the timing model until all phase jumps are removed and the solution is considered phase-connected. This means that all NS rotations are accounted for between observations. Sometimes DRACULA will give two or three valid solutions for partially phase-connected TOAs and it is always worth exploring all cases. In the case of all pulsars in this work, DRACULA output only one timing solution when all TOAs were phase-connected.

Once phase connection was achieved for each pulsar using TOAs spread in both frequency and time (roughly $2-4$ subbands and subintegrations for each observation), TOAs made from the refolded data were fully scrunched in time and split into usually two, but at most four, subbands each representing ~ 400 or 200 MHz bandwidths respectively. This was done in order to get the most precise astrometric parameter measurements, especially in the case of the short period binaries. Subbands with a center frequency of below 1100 MHz were omitted in all cases due to reciever cut-off; although this was only an issue on days split into four subbands. For PSR J1936+1805, TOAs with uncertainties greater than 100 μs were omitted from further analysis however, in almost all cases these omissions coincided with TOAs centered below 1100 MHz. In the few instances that a TOA centered at 1680 MHz was omitted due to the TOA uncertainty cut-off, the profile was checked manually in order to ensure a poor-quality signal. In all cases, high-frequency omitted TOAs displayed faint to no signal.

After the undergoing the refolding process in order to improve the profile (and therefore parameter) resolution, 76 phase connected TOAs were calculated for PSR J1851+0010, 52 for PSR J1853+0008, 164 for PSR J1936+1805 and 55 for PSR J1936+2142. These TOAs were timed in TEMPO2 using a least-squares fit to calculate the parameter uncertainties. The Solar system and clock corrections used were, respectively, the JPL's DE438 (Folkner and Park, 2018) and TT(BIPM19) (Guinot, 1988). Where the standard clock correction could not be used, clock corrections were made in accordance with BIPM guidelines[2] post-fit. Residual plots for all four pulsars are given in Figure 4.5 with Tables 4.1 and 4.2 giving the timing parameters for the best timing solution in each case.

The mean pulsar fluxes, as well as various pulse profile properties are tabulated in Table 4.3. The parameters probed were the pulse widths at 10 and 50% peak amplitude, w_{10} and w_{50}; the mean 1400 MHz phase-averaged flux density, S_{1400}, averaged from all profiles

[2]See ftp://ftp2.bipm.org/pub/tai/ttbipm/TTBIPM.2019

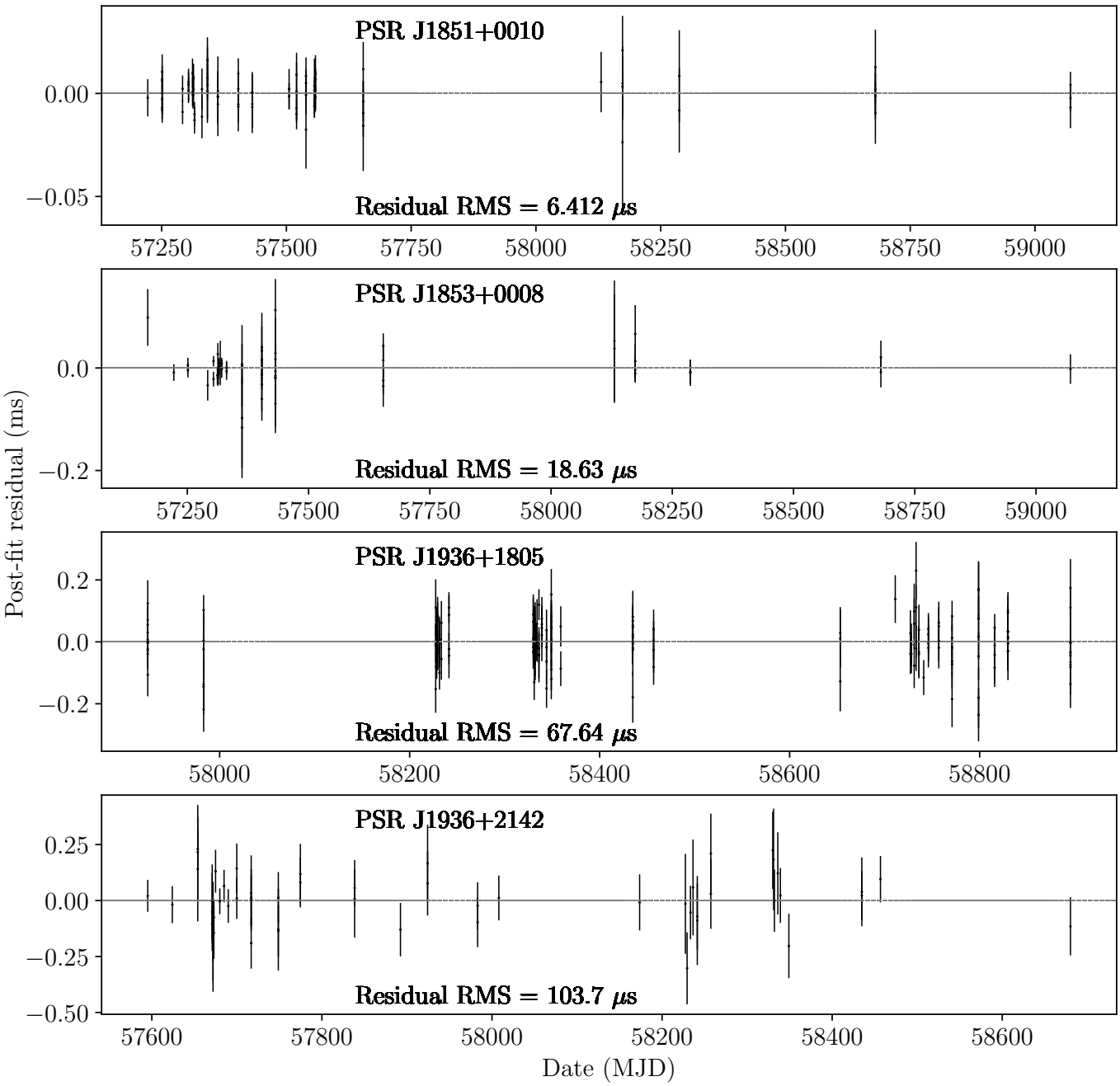


Figure 4.5: Post-fit timing residuals, in milliseconds, as a function of MJD for PSRs J1851+0010, J1853+0008, J1936+1805 and J1936+2142 as reported by TEMPO2, corresponding to the best solution after refolding the data. All observations were taken at a centre frequency of 1400 MHz.

Table 4.1: Timing solution for PSRs J1851+0010 and J1853+0008.

Fit and data-set		
Pulsar	PSR J1851+0010	PSR J1853+0008
Date range (MJD)	57222.1 – 59071.1	57168.3 – 59071.1
Data span (yr)	5.06	5.21
Number of TOAs	76	52
Solar System ephemeris	DE438	
Clock correction procedure	TT(BIPM19)	
Reference timing epoch (MJD)	57313.0	57924.0
Binary model	DDGR	
RMS timing residual (μs)	6.412	18.63
Reduced-χ^2	0.79	1.07
Observed quantities		
Right ascension, $\alpha_{\rm J2000}$	$18^{\rm h}51^{\rm m}02\overset{\rm s}{.}73753(8)$	$18^{\rm h}53^{\rm m}40\overset{\rm s}{.}4560(2)$
Declination, $\delta_{\rm J2000}$	$+00°10'15\overset{''}{.}721(3)$	$+00°08'03\overset{''}{.}915(8)$
Rotation frequency, ν ($\rm s^{-1}$)	43.785206576441(12)	29.93763478447(5)
First derivative of rotation frequency, $\dot{\nu}$ ($\rm s^{-2}$)	$-2.3026(14) \times 10^{-16}$	$-6.635(8) \times 10^{-16}$
Dispersion measure, DM ($\rm cm^{-3}pc$)	107.6195(17)	192.242(6)
$\dot{\rm DM}$ ($\rm cm^{-3}$ pc $\rm yr^{-1}$)	−0.0047(16)	–
Binary period, $P_{\rm b}$ (d)	2.003462167(11)	9.6129409(3)
Orbital eccentricity, e	0.0820775(5)	0.1601099(5)
Projected semi-major axis of orbit, x (lt-s)	9.177168(9)	26.582298(6)
Longitude of periastron, ω (deg)	298.3296(5)	160.9474(9)
Epoch of periastron, T_0 (MJD)	57234.858566(3)	57920.83266(2)
Total system mass, $M_{\rm tot}$ ($M_\odot$)	2.588(11)	2.6(2)
Companion mass, $m_{\rm c}$ ($M_\odot$)	1.15(2)	1.3(6)
Derived quantities		
Rotation period, P (ms)	22.838764007066(6)	33.4027723699(10)
First derivative of rotation period, $\dot{P}$	$1.2010(7) \times 10^{-19}$	$7.403(15) \times 10^{-19}$
Galactic longitude, ℓ	$33\overset{\circ}{.}03942(8)$	$33\overset{\circ}{.}3064(3)$
Galactic latitude, b	$0\overset{\circ}{.}165(3)$	$-0\overset{\circ}{.}44(14)$
NE2001 DM-derived distance (kpc)	3.0(9)	5.2(1.6)
YMW16 DM-derived distance (kpc)	3.0(6)	3.9(8)
Height above Galactic plane, z (kpc)	0.0086(17)	−0.029(10)
Characteristic age, $\tau_{\rm c}$ (Gyr)	3	0.7
Spin-down luminosity, $\dot{E}^\dagger$ ($10^{32}\,\rm erg\,s^{-1}$)	−3.98	−7.84
Surface magnetic field strength, $B_{\rm s}$ (10^9 G)	1.68	5.03
Advance of periastron, $\dot{\omega}^*$ ($\rm deg\,yr^{-1}$)	0.1186(4)	0.0081(5)
Mass function, f ($M_\odot$)	0.2067502(8)	0.21835(18)
Pulsar mass, $m_{\rm c}$ ($M_\odot$)	1.43(2)	1.3(6)

Figures in parentheses indicate the 1σ (68%) uncertainties in the least significant digits.
†Calculated using Equation 1.5 with $I = 10^{45}$ g cm^2.
*Measured using the DDFWHE (DDH) Shapiro-delay model in TEMPO.

Table 4.2: Timing solution for PSRs J1936+1805 and J1936+2142.

Fit and data-set		
Pulsar	PSR J1936+1805	PSR J1936+2142
Date range (MJD)	57924.3 − 58895.6	57594.2 − 58681.2
Data span (yr)	2.66	2.97
Number of TOAs	164	55
Solar System ephemeris	DE438	
Clock correction procedure	TT(BIPM19)	
Reference timing epoch (MJD)	57924.0	58137.0
Binary model	–	ELL1
RMS timing residual (μs)	67.636	103.66
Reduced-χ^2	1.15	1.02
Observed quantities		
Right ascension, $\alpha_{\rm J2000}$	$19^{\rm h}36^{\rm m}05\overset{\rm s}{.}6040(3)$	$19^{\rm h}36^{\rm m}54\overset{\rm s}{.}2712(10)$
Declination, $\delta_{\rm J2000}$	$+18°05'35\overset{''}{.}634(7)$	$+21°42'19\overset{''}{.}37(2)$
Rotation frequency, ν (s^{-1})	17.139401819995(14)	31.65670154435(4)
First derivative of rotation frequency, $\dot{\nu}$ (s^{-2})	$-2.75(3) \times 10^{-17}$	$-4.6(3) \times 10^{-17}$
Dispersion measure, DM ($cm^{-3}pc$)	126.129(9)	74.90(3)
Binary period, $P_{\rm b}$ (d)	–	0.757222866(3)
1st Laplace-Lagrange parameter, η	–	0.000028(16)
2nd Laplace-Lagrange parameter, κ	–	0.000009(16)
Projected semi-major axis of orbit, x (lt-s)	–	3.13795(2)
Epoch of ascending node, $T_{\rm asc}$ (MJD)	–	58136.5277278(15)
Derived quantities		
Rotation period, P (ms)	58.34509339955(5)	31.5888880147147(4)
First derivative of rotation period, $\dot{P}$	$9.36(10) \times 10^{-20}$	$4.6(3) \times 10^{-20}$
Galactic longitude, ℓ	$54\overset{\circ}{.}0571(3)$	57.3
Galactic latitude, b	$-1\overset{\circ}{.}271(7)$	0.354
NE2001 DM-derived distance (kpc)	4.7(1.4)	3.6(1.1)
YMW16 DM-derived distance (kpc)	3.6(7)	3.0(6)
Height above Galactic plane, z (kpc)	−0.07(2)	0.018(4)
Characteristic age, $\tau_{\rm c}$ (Gyr)	9.9	11
Spin-down luminosity, $\dot{E}^*$ (10^{30} erg s^{-1})	−19	−0.0000005
Surface magnetic field strength, $B_{\rm s}$ (10^9 G)	2.36	1.21
Orbital eccentricity, e	–	0.000030(15)
Longitude of periastron, ω (deg)	–	71.2(30.0)
Epoch of periastron, T_0 (MJD)	–	58136.68(6)
Mass function, f ($M_\odot$)	–	0.0578592(13)

Figures in parentheses indicate the 1σ (68%) uncertainties in the least significant digits.
*Calculated using Equation 1.5 with $I = 10^{45}$ g cm^2.

where both polarisation and flux calibration have been applied; the YMW16 distance (Yao et al., 2017) and the 1400 MHz luminosity, L_{1400}, approximated as $S_{1400}d^2$. The widths and fluxes were calculated using the `pdv` program in PSRCHIVE (Hotan et al., 2004). Pulse widths for PSR J1853+0008 and PSR J1936+1805 are discussed further in Section 4.6.

Table 4.3: Mean single pulse properties for the four pulsars taken across the date range in Tables 4.1 and 4.2.

PSR	w_{10} (ms)	w_{50} (ms)	S_{1400} (mJy)	d^* (kpc)	$L^\dagger_{1400}$ (mJy kpc^2)
J1851+0010	0.63	0.36	0.14	3.0(6)	1.3(3)
J1853+0008	2.6	1.0	0.21	3.9(8)	3.2(6)
J1936+1805$^\times$	2.3	1.4	0.53	3.6(7)	6.9(1.4)
J1936+2142$^\times$	4.0	2.0	0.09	3.0(6)	0.81(16)

*Derived from the YMW16 Galactic free electron distribution model with a 1σ (68% confidence) statistical error of 20%.
†Calculated from the approximation: $L_{1400} \approx S_{1400}d^2$. Uncertainties in L_{1400} are propagated from the distance under the assumption that the uncertainty in distance far exceeds any uncertainty in the flux.
$^\times$Profiles were divided through the off-pulse RMS for each hand of polarisation to obtain the quoted flux values (see Section 2.3 for an explaination).

4.4.1 Binary fitting

TOAs used for timing are input in the topocentric (telescope) frame of reference. This allows the user to convert to more useful frames of reference dynamically during the timing procedure instead of relying on a static, pre-determined barycentric conversion of the TOAs. However, for the purposes of probing NS spin parameters, barycentric folding periods can be made using Solar System distance estimates and geometry as given by a Solar system ephemeris such as DE438 (Folkner and Park, 2018).

In order to determine any possible binary character for the initial solutions to the PALFA pulsars, manually-determined barycentric folding periods were used to fit a general sinusoid (Equation 4.1) to the incoherent barycentric folding periods.

$$f(t) = t_0 \sin\left(\frac{2\pi t}{P_\mathrm{b}} + a\right) + \langle P \rangle \tag{4.1}$$

where t_0 is the first epoch of the data set, a is the pulse phase, found through a least-squares fit on the folding spin periods, and $\langle P \rangle$ is the mean spin period across all observations. This initial guess produces a binary solution for a circular Keplerian orbit with $e = \omega = 0$ and the other three parameters: T_0, x and P_b non-zero, where the initial guesses for T_0 and P_b were made using a range of values spanning a year (from MJD $57100 - 57465$) for T_0 and from $0.1 - 10$ days for P_b. The initial guess for x was obtained via Kepler's third law and the guess for P_b. These guesses were made until the fit sinusoid produced the lowest χ^2.

To obtain the full binary solutions, a grid was created in $e - \omega$ space with e running

from 0.0 to 1.0 in 0.00001 increments and ω running from 0 to 360deg in 2.5deg increments for 14,400,000 total iterations. Each cell of the grid was combined with the three binary parameters in the circular solution and attached to a temporary ephemeris file, which was then analysed through TEMPO using the incoherent (not phase-connected) TOA file. Any fits that resulted in a reduced-$\chi^2 < 2$ were investigated further by attempting phase connection in TEMPO with the Blandford-Teukolsky (BT, Blandford and Teukolsky, 1976) or, in the case of PSR J1936+2142, the ELL1[3] (Lange et al., 2001) binary model, until a solution was found that minimised both the residual RMS and the reduced-χ^2. In the case of PSR J1936+2142, the final phase-connected solution was achieved through use of the phase-connecting program DRACULA (Freire and Ridolfi, 2018), which can accurately determine the number of rotations from one TOA to the next by adding unit rotations, N, between incoherent TOAs until the value of N produces the lowest residual χ^2.

The final phase connected solution was tested against a sample of 10 points, drawn from a Gaussian distribution with the folding barycentric rotation periods as the mean and the 2σ Gaussian width as the uncertainty. For an example of this, see Figure 4.6, which shows that the solution describes each period within that variance.

Once refolded with the intial solution, the TOAs for PSR J1851+0010 and PSR J1853+0008 were fit to the DDGR (Taylor, 1987) binary model in order to find component masses, and then to the DDH (Freire and Wex, 2010) model in order to find the extent of Shapiro delay contributing to those mass constraints. These models are discussed in Section 3.3.

[3]ELL1 uses the 1st and 2nd Laplace-Lagrange parameters which are coefficients of the small-eccentricity first order Roemer delay expansion and are orthogonal projections of the eccentricity ($\eta \equiv e \sin\omega$ and $\kappa \equiv e\cos\omega$) and have the relation $\eta^2 + \kappa^2 = e^2$, and also parametrises the epoch of the ascending node, $T_{\rm asc}$, as opposed to the epoch of periastron.

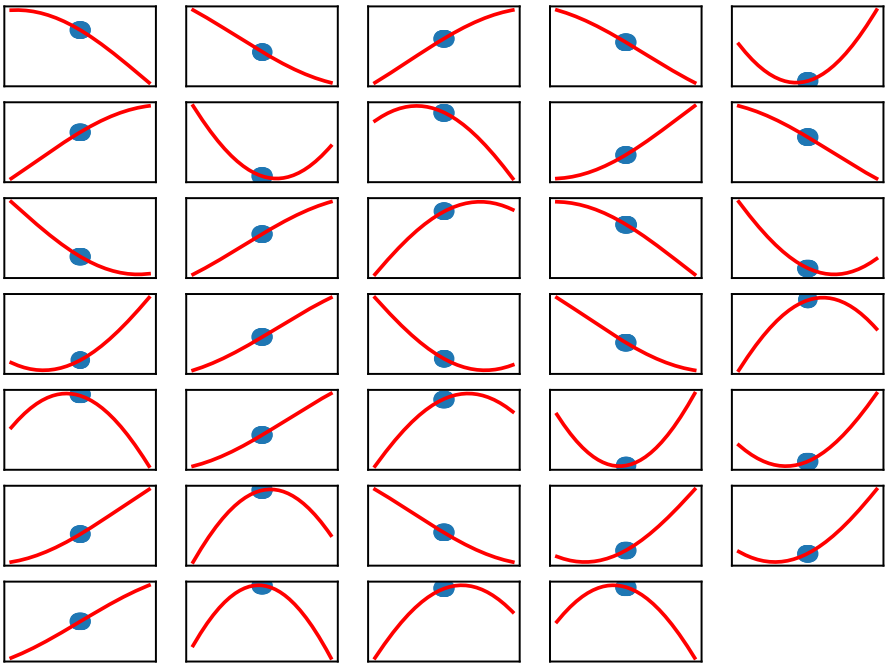


Figure 4.6: A local point zoom of the full orbital fit for PSR J1936+2142. Each panel represents one observing epoch. In each panel, ten points (blue) centered around the local folding period were sampled. The red solid line represents the best binary fit obtained via `tempo`.

4.5 Results and discussions

As a result of the timing analyses administered on these four pulsars, the $P-\dot{P}$ diagram has been updated to reflect the derived periods and spin-down rates (Figure 4.7), as determined by TEMPO2. Generally, all four pulsars appear to lie in the $P-\dot{P}$ region consistent with both the DNSs and the growing population of intermediate mass binary pulsars (IMBPs), binaries with high-mass WD companions (e.g. Ferdman et al., 2010). In all cases, their spin periods suggest some degree of recycling; this is to be expected for PSRs J1851+0010, J1853+0008 and J1936+2142 as their best-fit timing solutions account for a binary orbit, however PSR J1936+1805 shows no such binary trends in the best-fit residuals, despite ample data for such a trend (see Section 4.5.3). This region of the $P-\dot{P}$ diagram also implies the ages of all the pulsars are on the order of 1 Gyr – consistent with the ages of known recycled pulsars and the timescales in which recycling needs to occur (see e.g. Tauris et al., 2017). For the three binary systems, their position on the $P_\mathrm{b}-P$ diagram is given in Figure 4.8, which shows them to be typical of other partially-recycled binary pulsar systems.

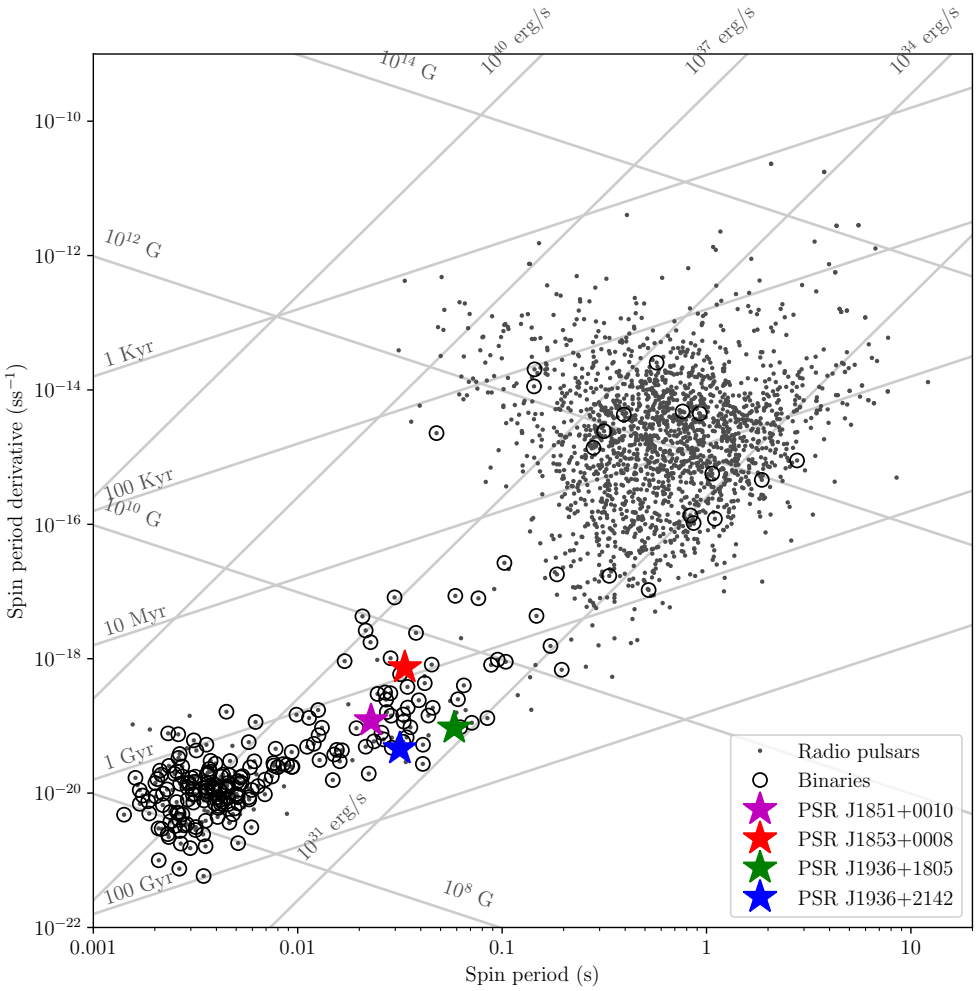


Figure 4.7: The $P-\dot{P}$ diagram, including the four pulsars: PSRs J1851+0010, J1853+0008, J1936+1805 and J1936+2142.

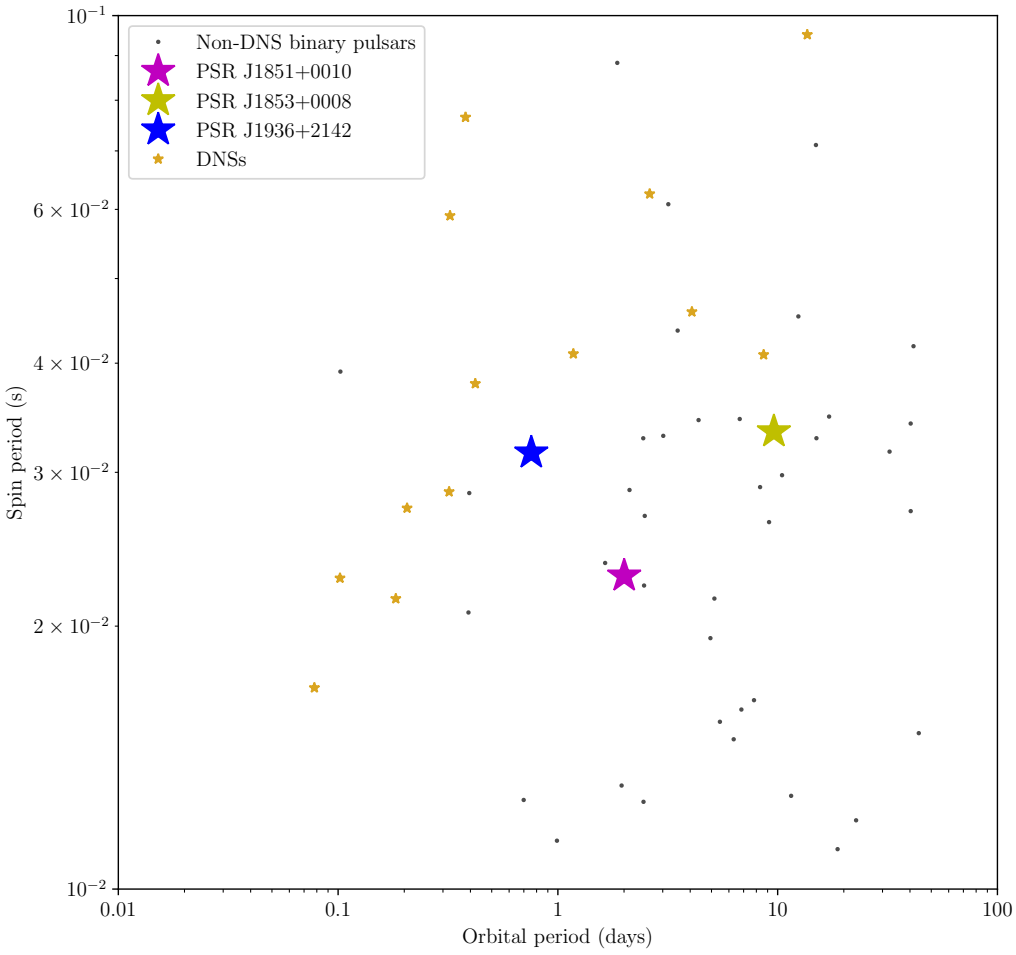


Figure 4.8: The $P_{\rm b} - P$ diagram outlining the three binary pulsars: J1851+0010, J1853+0008 and J1936+2142. The gold stars represent the DNS population as given in Table 3.4. Note that the orbital period is on a logarithmic scale whereas the spin period is on a linear scale.

4.5.1 PSR J1851+0010

PSR J1851+0010 has been discovered to be a member of a relatively compact ($x \approx 9.17\,\mathrm{lt\text{-}s}$), moderately circular ($e \approx 0.08$) binary of orbital period 2 d (48.1 hr). Neither the measured Keplerian orbital parameter, nor the pulsar's 22.8 ms spin period, give away the exact nature of this binary with absolute certainty. On the one hand, the eccentricity is high for a low-mass NS-WD system, but on the low end of expected DNS eccentricities. Given its compact orbit, it more likely falls under the umbrella of the DNSs although some IMBPs have compact orbits. The eccentricity is however notably similar to some DNSs; both the spin period and eccentricity are notably similar to several DNS systems (e.g. PSR J0737−3039 ($e \approx 0.088$; Kramer et al., 2006) and PSR J1756−2251 ($e \approx 0.18$; Ferdman et al., 2014)).

Timing has yielded a significant value for the relativistic advance of periastron as $0.1185 \pm 0.0008\,\mathrm{deg\,yr^{-1}}$. According to GR (Equation 1.28), this gives a precise total mass measurement of $2.58 \pm 0.03\,M_\odot$, which is more indicative of a DNS system than an IMBP although some NS-WD systems have been found in this total mass region (e.g. PSR J1614−2230; Arzoumanian et al., 2018). To determine the component masses of the system, another PK-parameter must be measured. Assuming that the orbit is not highly inclined, the obvious candidate to try will be γ. If the first-born NS (usually the observed pulsar) mass is assumed to be the "canonical" value of $1.4\,M_\odot$ (Thorsett and Chakrabarty, 1999), γ is expected to be 0.90 ms (from Equation 1.29). Although, if this system is indeed a DNS system, a more accurate value for γ may be given by assuming $m_\mathrm{p} = 1.33 \pm 0.09\,M_\odot$ (Özel and Freire, 2016). In this case, γ is expected to be 0.97 ms, with a simulated uncertainty of $\pm 74\,\mu$s.

The DDGR binary model gives precise component mass boundaries of $1.43 \pm 0.04\,M_\odot$ and $1.15 \pm 0.04\,M_\odot$ for the pulsar[4] and its companion, respectively. If these mass constraints are to be believed, this would make the companion to PSR J1851+0010 the lightest NS currently known (cf. Martinez et al., 2015), although this low companion mass does not require a reconsideration of the proposed evolution pathway, which is likely heavy tidal stripping followed by an iron core-collapse supernova (see below and also Chapter 3). Using the DDH Shapiro delay timing model, the orthometric h_3 parameter was found to be $2.70 \pm 0.19\,\mu$s. The masses derived from this model adequately explain the component masses as determined by the DDGR model and both models are presented in Figure 4.9, which contains overlayed mass-mass contour plots, calculated using Bayesian joint posterior probability density theory (see Splaver et al., 2002, for more information and Section 3.4.2 in this work for the method of implementation), for both the DDH and DDGR models.

The period spin-down rate, $\dot{P}$, was found to be $(1.2010 \pm 0.0014) \times 10^{-19}\,\mathrm{s\,s^{-1}}$, placing the pulsar in the region of the $P - \dot{P}$ diagram containing the IMBPs and the DNSs. This $\dot{P}$ cannot yet be interpretted as the true spin-down rate of the NS, as no significant proper motion has been observed. Therefore, any kinematic effects on $\dot{P}$ could not be accounted for. Assuming that the measured value for $\dot{P}$ is entirely due to the proper motion (Shklovskii, 1970) so that the intrinsic spin period, $\dot{P}_\mathrm{int} = 0$, Equation 3.10 places

[4]Pulsar mass is obtained from $M_\mathrm{tot} - m_\mathrm{c}$, which is allowed if $\sigma_{M_\mathrm{tot}} << \sigma_{m_\mathrm{c}}$ so that $\sigma_{m_\mathrm{c}} \sim \sigma_{m_\mathrm{p}}$.

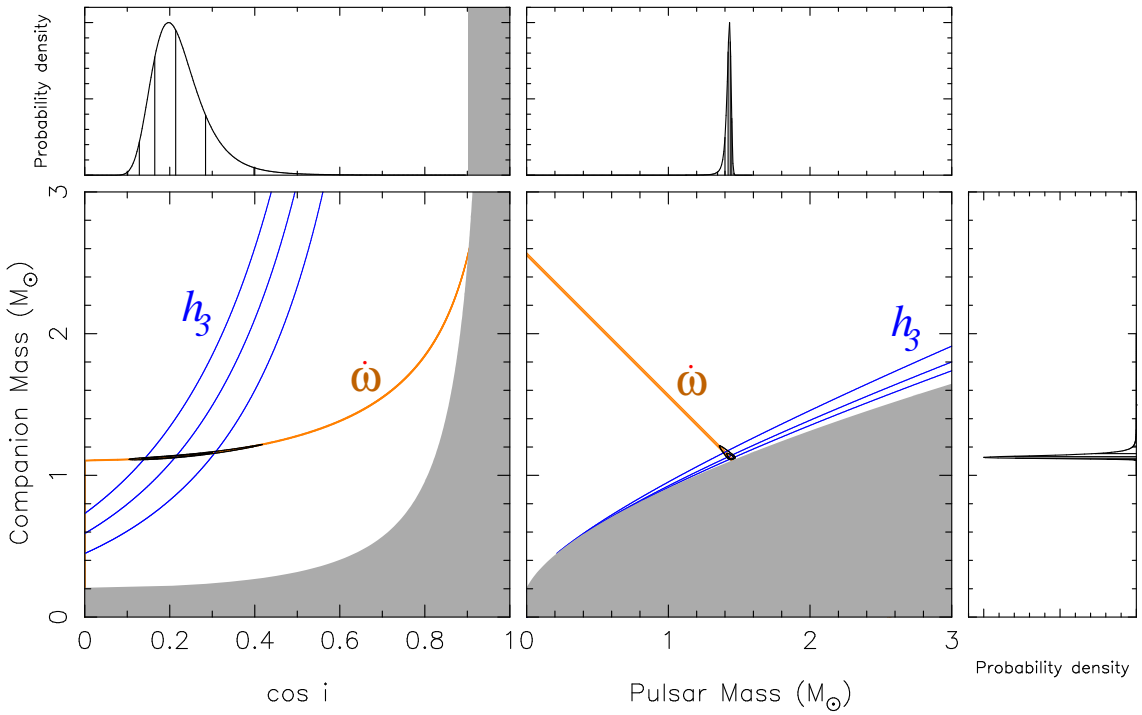


Figure 4.9: Mass diagrams for PSR J1851+0010 showing the GR-derived mass constraint contours (68%, 95% and 98.5% confidence levels) together with $\dot{\omega}$ and h_3, fit in the DDH model, with companion mass plotted against $\cos i$ (left panels) and pulsar mass (right panels). The orange region is $\dot{\omega}$ and the blue region is h_3, as reported by TEMPO2. The gray area represents $\sin i > 1$. The three plots on the outside are the three 2D probability density functions for each set of variables.

an upper limit of $27\,\mathrm{mas\,yr^{-1}}$ on the total proper motion for this system. Using the median YMW16 DM-derived distance of 3.0 kpc, Equation 3.8 gives an upper bound to the transverse velocity of $\sim 380\,\mathrm{km\,s^{-1}}$, corresponding to a Local Standard of Rest velocity, v^{LSR}, of $\sim 350\,\mathrm{km\,s^{-1}}$; this is concurrant with high-kick SN models, although it is ofcourse very unlikely that all of the observed period derivative is due to proper motion, so this velocity is likely to be a vast overestimate.

Taking a more conservative estimate for the proper motion ($\mu_{\mathrm{tot}} \sim 10\,\mathrm{mas\,yr^{-1}}$) and also taking into account the measured total mass, eccentricity and spin properties, there is a high liklihood that this system is a DNS. If the proper motion is measured to be small ($\lesssim 6\,\mathrm{mas\,yr^{-1}}$), this would place it firmly in the group of DNS systems that probably evolved due to ultra-stripped iron core-collapse (FeCCSN; Tauris et al., 2017) of the He-shell during the second supernova, similar to e.g. PSRs J0737−3039 (Kramer et al., 2006; Ferdman et al., 2013), J1756−2251 (Ferdman et al., 2014) and J1829+2456 (Haniewicz et al., 2021). If the proper motion is found to be large, the binary will share some properties, and therefore possibly some evolutionary links, with e.g. PSRs B1534+12 (Fonseca et al., 2014) and possibly J1913+1102[5] (Ferdman et al., 2020). At the current level of precision, now that the component masses have been measured, TOA simulations made with the LIBSTEMPO Python package[6] predict that two more years of four-weekly cadence timing

[5] Although, the evolution of PSR J1913+1102 may actually be symmetric.
[6] http://vallis.github.io/libstempo/

data will produce proper motion measurements with a 95% confidence, which will allow a more precise estimate of the space velocity.

The first derivative to the DM has also been measured to be (-0.0047 ± 0.0032) cm^{-3} pc yr^{-1}, which implies that the dispersion is weakening over time. If this is true and the second derivatives are negligible, then the system could be moving towards us in space due to the second supernova kick. However, DNS systems with these orbital parameters tend to have small predicted natal kicks (Tauris et al., 2017) so it is unlikely that a DNS would shift the DM by this much per year even if the direction of motion was completely radial. Another possible situation is that our relative position to PSR J1851+0010 in the galaxy is changing our line of sight significantly to present a noticable change in DM over time.

An extremely light neutron star

The mass for the companion to PSR J1851+0010 measures at $1.15 \pm 0.04\,M_\odot$. As mentioned, this could entail the lightest NS known to date. The first argument one may suggest is that this companion is a massive WD, however owing to the orbital eccentricity and binary separation, no NS-WD evolutionary channels fit whereas DNS evolutionary channels do. A NS this light could have potentially evolved from zero-age main sequence stars of very asymmetric mass. A comparison can be made to the zero-age evolution of PSR J0751+1807 (Fortin et al., 2016) in which it is conjectured that the initial main sequence binary has masses of $15\,M_\odot$ and $1.6\,M_\odot$. Following a CE phase and tidal stripping, the lower mass NS could have been born with a mass as low as $1.05 - 1.30\,M_\odot$ when taking into account accretion during spin-up. In the case of PSR J1851+0010, the companion NS must have been born with a similar mass to this or a mass slightly below $1\,M_\odot$.

A metallicity (post-He) study on PSR J1851+0010 and its surroundings could support these claims. Metallicity is proportional to opacity and high opacity is indicative of a high mass-loss period (for example, very tidal stripping) (see e.g. Pejcha and Thompson, 2015, and references therein). Also compare to the low-mass companion of PSR J1411+2551, supernova simulations of which have shown it to have possibly been born from a main sequence star of $\approx 2.5\,M_\odot$ and a low kick (Martinez et al., 2017); akin to the discussion above for PSR J1851+0010. Follow up timing should concentrate on constraining the masses of PSR J1851+0010 further. Current NS equations-of-state predict spun-up NS masses in the range $1.1 - 1.8\,M_\odot$ so if the mass of the companion is lower than $1.1\,M_\odot$, this could rule out many equations-of-state.

4.5.2 PSR J1853+0008

PSR J1853+0008 has been found to be a member of a wide ($P_\mathrm{b} = 9.61\,\mathrm{d}$) binary with moderately high eccentricity ($e \approx 0.16$). Although high eccentricity is often indicative of two compact bodies (i.e. a DNS), such a wide orbit in relation to this eccentricity may be more suggestive of an NS-WD binary with a high-mass WD companion (see analyses by Tauris et al., 2013). The advance of periastron for the system has been measured to be $0.0081 \pm 0.001\,\mathrm{deg\,yr^{-1}}$, which gives a GR-derived total mass of $2.6 \pm 0.4\,M_\odot$. Given the uncertainity in this mass measurement, it is not yet possible to conclude whether the

system is either a DNS or an IMBP (or possibly a NS-main sequence binary, although there is no known evidence of any potential optical counterpart). However, the relatively high eccentricity coupled with the observed 33.4-ms spin period make this system a strong DNS candidate. More confirmation is provided by the minimum possible mass, derived to be $\sim 1\,M_\odot$ from the orbital constraints – high for a WD, although still under the Chandrasekhar mass of $1.4\,M_\odot$.

The DDGR fit for companion mass and total mass gives rather wide 95% confidence component mass constraints, with a total mass of $2.6 \pm 0.4\,M_\odot$ and a companion mass of $1.3 \pm 1.1\,M_\odot$ inferring a pulsar mass of $1.3 \pm 1.1\,M_\odot$. These mass constraints cannot be used to make any definitive predictions on evolution or on the nature of the companion. In order to determine a possible explaination for these mass constraints, the DDH model was applied to the TOAs and the value for h_3 was found to have the expected nominal value ($\sim 1\,\mu$s), with a 2σ upper bound of approximately $8\,\mu$s. More data will be needed in order to disentangle parameters within the models.

The YMW16 DM-derived distance model gives $d = 3.9 \pm 0.8\,$kpc (68%), meaning that any space velocity estimates made will be quite broad (as velocity estimate errors, as determined by Monte-Carlo simulations, scale with the error in the distance (see Section 3.4.3)). The proper motion for this system is still not directly measurable, but the upper limit, again assuming a maximum Shklovskii contribution (Shklovskii, 1970) is calculated to be $48\,\mathrm{mas\,yr^{-1}}$. This leads to an upper bound estimate on the transverse velocity of $\sim 890\,\mathrm{km\,s^{-1}}$ and on v^{LSR} of $\sim 800\,\mathrm{km\,s^{-1}}$, which implies a large natal kick most likely imparted through a second supernova event, if the system is indeed a DNS. This maximum velocity estimate cannot be accounted for from the first supernova alone and it is clear that a good deal of mass transfer recycling occurred at the X-ray binary accretion stage.

Future timing campaigns for this pulsar should focus on constraining orbital parameters. Much like PSR J1851+0010, the current orbital coverage for PSR J1853+0008 is poor due to the sparseness of the dataset and the manner in which most of the data was taken[7]. A more precise value for $\dot{\omega}$ as well as further PK-parameters is essential for determining the nature of this binary system and it is predicted that 10% precision on $\dot{\omega}$ will be possible with $1-2$ more years of timing. It is unlikely that any other PK-parameters will be measureable for some time unless the inclination of the orbit happens to be advantageous ($i \approx 90$deg) due to the width of the orbit. Applying the orthometric Shapiro delay parametrisation (DDH) to the DD model did not give statistically significant results; even after refolding. $\dot{P}_\mathrm{b}$ is not expected to be measureable, even after five years of timing without the use of possibly multiple dense campaigns to achieve near-full orbital coverage. The upper bound on γ is derived to be $\approx 9\,$ms under GR and the median delay effect due to γ is expected to be 3.4 ms.

[7] i.e. not at optimally devised times based on orbital calculations but rather during a routine sky scan

4.5.3 PSR J1936+1805

PSR J1936+1805 is an isolated pulsar with a high spin period (58.3 ms) relative to the normal pulsar population. It also has a very stable observed spin-down rate, $\dot{P}$, of $9.36 \pm 0.18 \times 10^{-20}$, implying that this pulsar is very old (~ 10 Gyr). These values imply this pulsar was partially recycled. At the current time, proper motion measurements cannot be made so corrections for kinematic effects cannot yet be made. Due to careful subbanding of the data, a precise value for the DM is determined to be $126.13 \pm 0.02\,\mathrm{cm}^{-3}\,\mathrm{pc}$, leading to a YMW16 DM-derived distance estimate of 3.6 ± 0.7 kpc (68%). Although this distance estimate has a somewhat large statistical uncertainty (20% at 1σ), there is little hope for an improvement to this, as a parallax measurement is unlikely for a system potentially so far away.

It is likely that this pulsar is a *disrupted recycled pulsar* (DRP): a pulsar that would have otherwise been the first-formed NS in a DNS, except that the resulting kick and mass-loss from the second supernova disrupted the system (e.g. Lorimer et al., 2004). If this is the case, then the observed spin parameters would fit well with other DRPs such as PSR J0609+2130 (Lorimer et al., 2004) and PSR J2053+1718 (Camilo et al., 1993; Brinkman et al., 2017).

There may be plausible alternative evolution pathways that could lead to the observed spin parameters. NSs observed in the center of supernova remnants are obviously young, however some still have relatively small ($\sim 10^9$ G) surface magnetic field strengths and large characteristic ages, similar to the observed DRP population, in a class of stellar remnant known as a central compact object (CCO; Halpern and Gotthelf, 2010). CCOs are characterised as strong thermal X-ray emitters and, so far, no confirmed DRP has been observed as an X-ray source (Gotthelf et al., 2013) nor have any radio pulsars been found as CCOs. This implies that CCOs and DRPs have evolved through different mechanisms despite having similar spin properties. Calculating limits on thermal emission for radio pulsars relies heavily on the NS equation-of-state and has not been explored in this work although a population study on DRP thermal emission is a proposed next step. Tight constraints on the proper motion of this system, from further timing, will help to determine the evolution scenario of PSR J1936+1805 more precisely using subsequent velocity estimates.

4.5.4 PSR J1936+2142

When investigating the barycentric folding periods of the incoherent data for PSR J1936+2142, the best fit (before any attempt at phase connection) occurs when accounting for a binary orbit with $P_\mathrm{b} \approx 0.757$ d. To ensure this was the best possible binary period fit, 99,900 equally spaced values for P_b ranging from $0.1 - 100$ days (0.001-d steps) were fit to the barycentric periods. This was obtained using a least-squares fit on the set of barycentred epochs and barycentric folding periods and by fitting these to a general sinusoid (4.1).

The reduced-χ^2 was calculated for each combination of spin period and modelled orbital period, and an orbital period of ≈ 0.757 d produced a χ^2_r closest to one (Figure 4.10). This P_b was then iteratively fit to a non-circular orbit as described in Section 4.4.1, providing excellent agreement at each epoch when compared with ten evenly sampled points centered

around the barycentric folding period (see Figure 4.6 and the explaination of this method in Section 4.4.1).

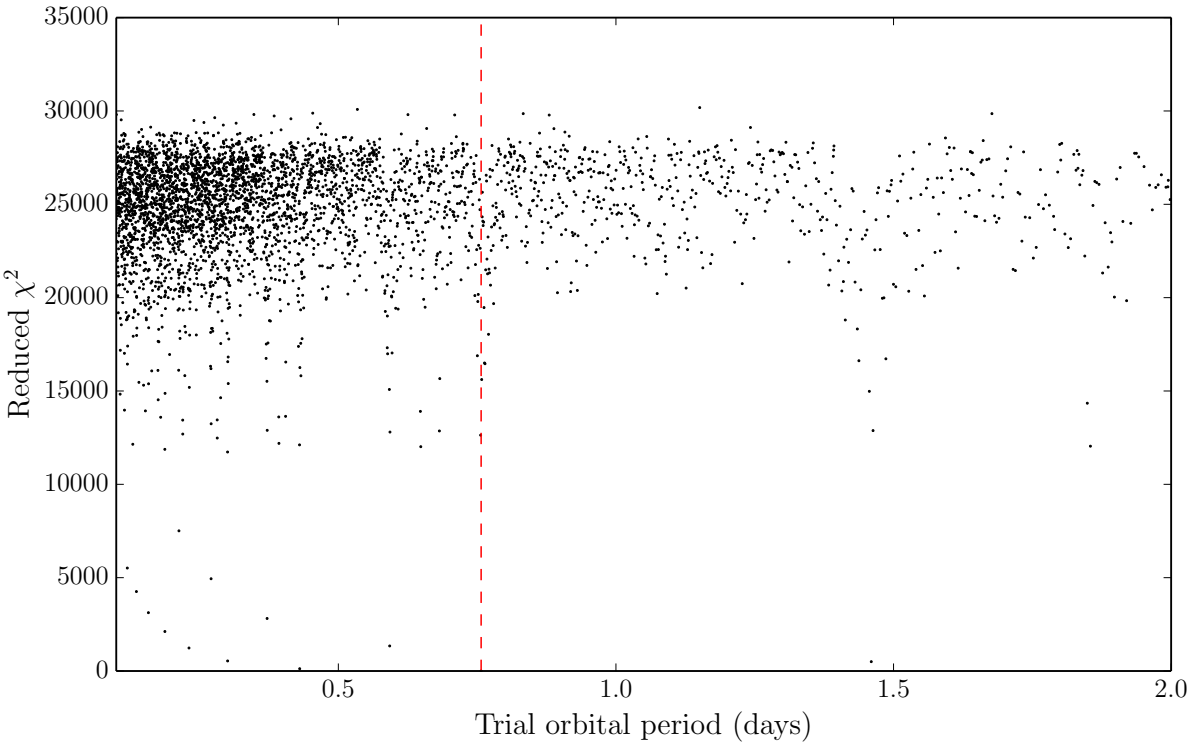


Figure 4.10: A periodogram displaying reduced-χ^2 with respect to the trial orbit periods for PSR J1936+2142 after having been fit to Equation 4.1. Orbital periods ranging from $0.1 - 100$ days were tested in 0.001 day increments. The red dashed line represents the P_b that produced the lowest χ^2_r (0.757 days). A harmonic can also be seen at double this orbital period.

When the binary orbit was confirmed, phase connection was attempted iteratively: first with the dense campaign TOAs in order to get a preliminary solution, and then to the rest of the data using DRACULA. The first binary model applied to the TOAs was the DD model, which gave a low eccentricity ($\sim 3 \times 10^{-5}$); the ELL1 binary model (Lange et al., 2001) was therefore used in order to potentially avoid covariances between T_0, ω and e. The ELL1 model reparametrises these three orbital quantities in terms of three new parameters: the epoch of the ascending node T_asc, and the first and second Laplace-Lagrange parameters, η and κ:

$$T_\mathrm{asc} \equiv T_0 - \frac{P_\mathrm{b}\omega}{2\pi} \tag{4.2}$$

$$\eta \equiv e \sin\omega \tag{4.3}$$

$$\kappa \equiv e \cos\omega \tag{4.4}$$

Under this transformation, the eccentricity was found to be 0.000029 ± 0.000016. The ELL1 results also give a derived value for ω as 71 ± 30 deg and T_0 as 58136.68 ± 0.06 MJD.

PSR J1936+2142 is in a near-circular orbit that has an old characteristic age ($\sim 11\,\mathrm{Gyr}$) – further evidence against it being a DNS system. Its circularity implies no second supernova having taken place and its age implies the second supernova should have happened by now if there was going to be one. The upper bound on the proper motion is found to be $14\,\mathrm{mas\,yr^{-1}}$ (95%) implying a maximum transverse velocity of $\sim 200\,\mathrm{km\,s^{-1}}$; this is in stark constrast to the other binary pulsars in this chapter; PSRs J1851+0010 and J1853+0008, which both have much larger maximum velocity upper limits than PSR J1936+2142. This may suggest that PSR J1936+2142 is a low mass binary pulsar (LMBP) or an IMBP, having had no second supernova, whereas the other two are likely DNSs.

It has been suggested that IMBPs ought to have lower Galactic heights compared with the LMBP population (see Camilo et al., 2001). This can plausibly be attributed to the relative difference in combined progenitor mass for these two types of system. An IMBP will likely have a smaller peculiar velocity than a LMBP assuming a similar kick force from their respective supernovae due to its larger total mass. PSR J1936+2142's height above the Galactic plane is $0.018 \pm 0.08\,\mathrm{kpc}$, which is low even for the observed IMBPs and much lower than the median scale height for the LMBPs. The transverse velocity estimate above is also indicative of a more massive progenitor than expected for LMBPs, adding to the evidence that this pulsar is a member of an IMBP. This observed Galactic scale height relationship may also be a selection bias due to recent pulsar searches being more sensitive to the near-Galactic-plane region, although this relationship has also been observed between the HMXBs and the LMXBs (van Paradijs and McClintock, 1995). PSR J1936+2142 has a DM-derived distance of $\approx 3\,\mathrm{kpc}$; similar to pulsars in known IMBP systems such as PSR J1802−2124 ($d = 2.94\,\mathrm{kpc}$; Ferdman et al., 2010) and PSR J1435−6100 ($d = 3.3\,\mathrm{kpc}$; Camilo et al., 2001; Manchester et al., 2001), so is expected to follow similar Galactic scale height, component mass and velocity trends as these pulsars.

Further data for PSR J1936+2142 may allow a significant of $\dot{\omega}$, although simulations do not support this. More timing may give reasonable bounds on the proper motion of this system and hence an estimate on the peculiar velocity, which will allow for tighter constraints on the evolution scenario of this system.

4.6 Profile analysis of PSRs J1853+0008 and J1936+1805

PSRs J1853+0008 and J1936+1805 show significant double peaks in their standard profile at all but the lowest of resolutions. It is therefore worth investigating any time-dependent variations in the integrated pulse profile and determine whether these variations lead to significant TOA inaccuracies, compared with a single standard profile for the enitre data set. In order to determine possible profile variations, the search-mode data for each pulsar were scrunched fully in time, frequency and polarisation to obtain an integrated profile. To set up a direct comparison between epochs, the profiles were: crudely flux calibrated by dividing them by their off-pulse RMS in each hand of polarisation; normalised, by dividing them by their maximum value; and finally, centred in phase. These integrated profiles were fit to two Gaussians to reflect the dual-peak pulse structure observed. The Gaussian-fit profiles are plotted as waterfall plots in Figures 4.11 and 4.12.

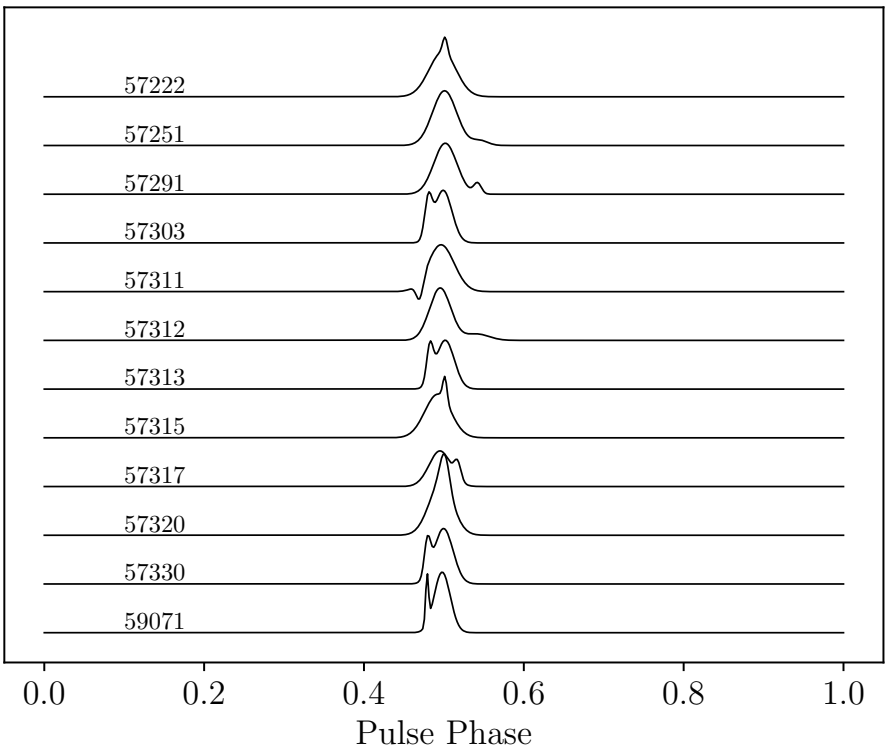


Figure 4.11: A waterfall plot displaying all the search-mode data for PSR J1853+0008 as Gaussian-smoothed integrated profiles. All profiles have been normalised with respect to their maximum value and vertically offset for visual clarity. As such, the y-axis is in arbitrary flux units.

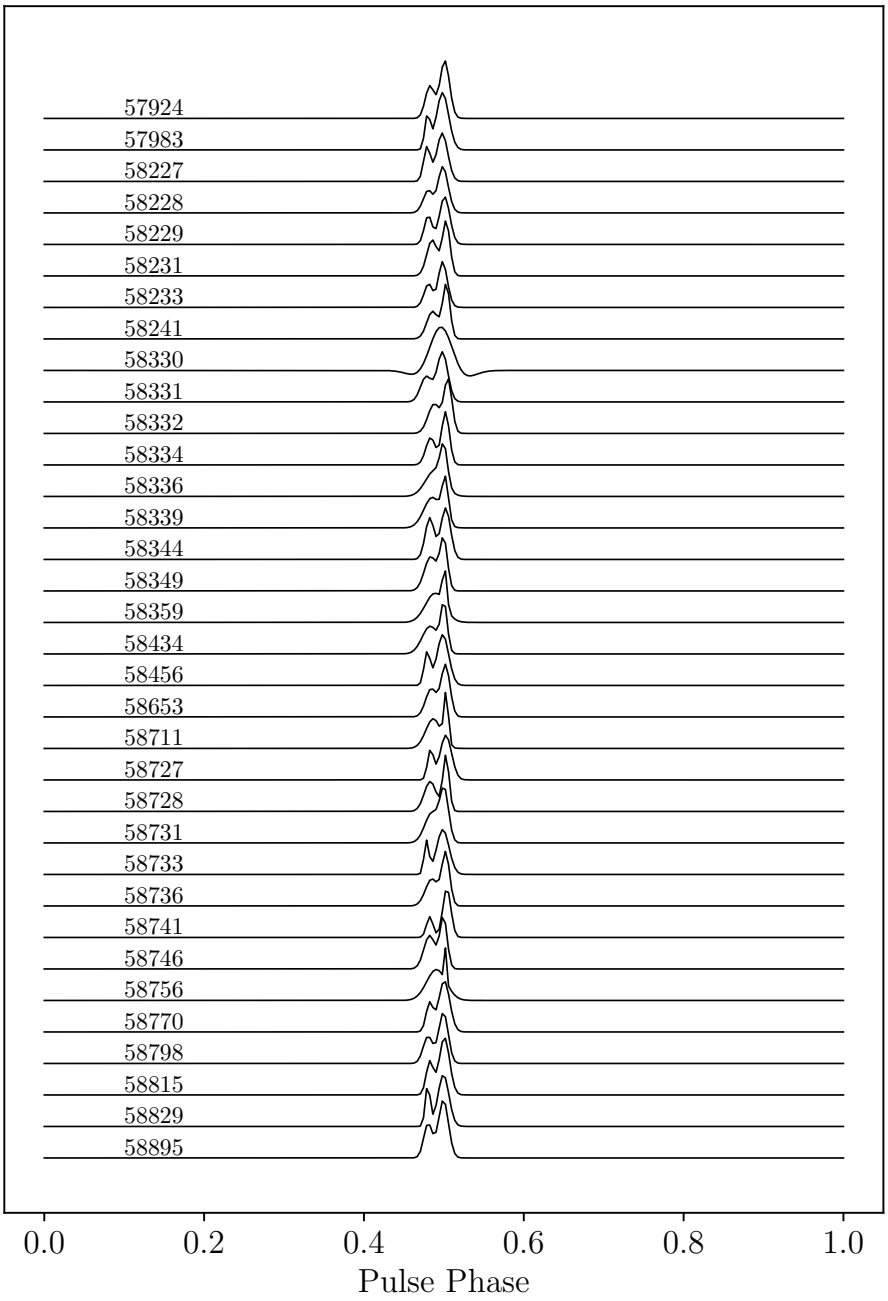
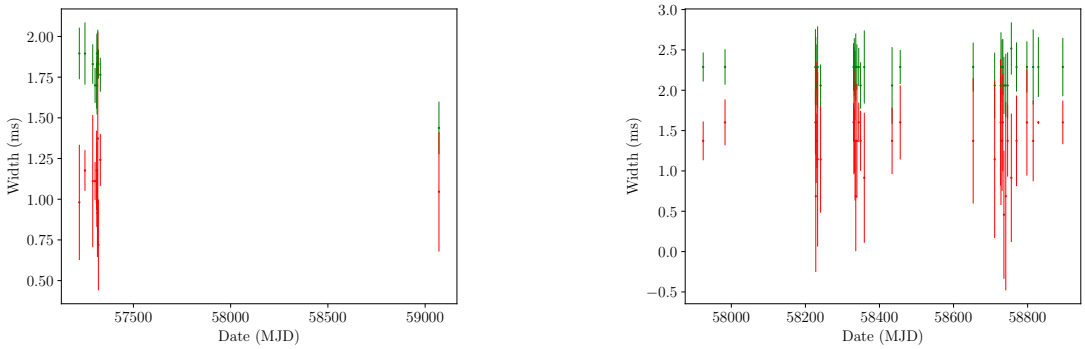


Figure 4.12: A waterfall plot displaying all the search-mode data for PSR J1936+1805 as Gaussian-smoothed integrated profiles. All profiles have been normalised with respect to their maximum value and vertically offset for visual clarity. As such, the y-axis is in arbitrary flux units.

In both cases, it is clear that the double-peak pulse structure exists and is not just a product of poor folding or RFI. In order to determine possible variation, w_{10} and w_{50} were calculated for each date as well as their ratio, with uncertainties determined via a bootstrap method (see Appendix A.4) with 10,000 iterations. A high w_{10} : w_{50} ratio implies a broad pulse whereas a ratio close to unity is suggestive of a narrow pulse. A significant change in these widths and their ratio over time may be due to gyroscopic precession, the conic precession of the rotation axis, and can potentially allow an estimate of the spin-emission misalignment angle under certain beam geometry theories (e.g. Perera et al., 2014). Investigation showed that profile shape change over time, determined by the ratio of widths, is consistent with zero (at a 95% confidence) for both PSR J1853+0008 and PSR J1936+1805 (Figures 4.13 and 4.14). This is to be expected for PSR J1853+0008 as the gyroscopic precession period is expected to be approximately 1400 centuries. The biggest issue comes from the fact that profiles from some days are significantly worse Gaussian fits than other dates (see e.g. 57315 for PSR J1853+0008), but this is accounted for in the uncertainty estimates.

In order to compare the integrated pulse profiles directly for PSR J1936+1805, the raw integrated (not Guassian fitted) profile for each day was subtracted from the integrated profile from 57924 (Figure 4.15), all in units of 'local maximum profile flux'. Difference profile RMS values close to zero confirm negligible change in the pulse profile and significant spikes in the difference profile, especially in the on-pulse region, imply profile shape variation as any noise contributions should average out to approximately zero.

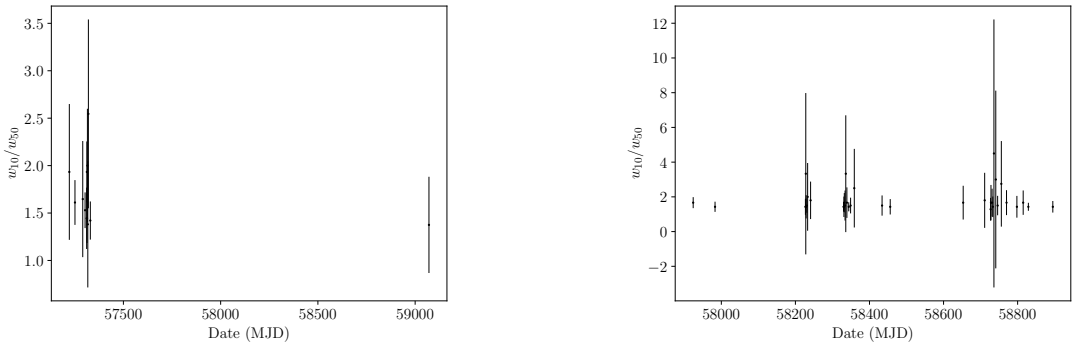


(a) PSR J1853+0008

(b) PSR J1936+1805

Figure 4.13: The time-evolution of pulse width, in milliseconds, at 10% maximum (green solid line) and 50% maximum (red dashed line) for PSR J1853+0008 (left) and PSR J1936+1805 (right). Data was taken from search-mode observations only.

Using an epoch-specific standard profile for each day of observation, instead of using one standard profile over the whole data sets, did not change TOA values by more than a few microseconds and TOA errors also improved by less than a few microseconds, confirming that the time evolution of the pulse shapes for both PSR J1853+0008 and PSR J1936+1805 is negligible. The difference plot also implies little variation in the pulse shape of PSR J1936+1805. This correlates well with the stable width ratios observed above. In the case of PSR J1853+0008, the time-stability of the pulse profile may hint



(a) PSR J1853+0008

(b) PSR J1936+1805

Figure 4.14: The time-evolution of pulse width ratios for PSR J1853+0008 (left) and PSR J1936+1805 (right). Data was taken from search-mode observations only.

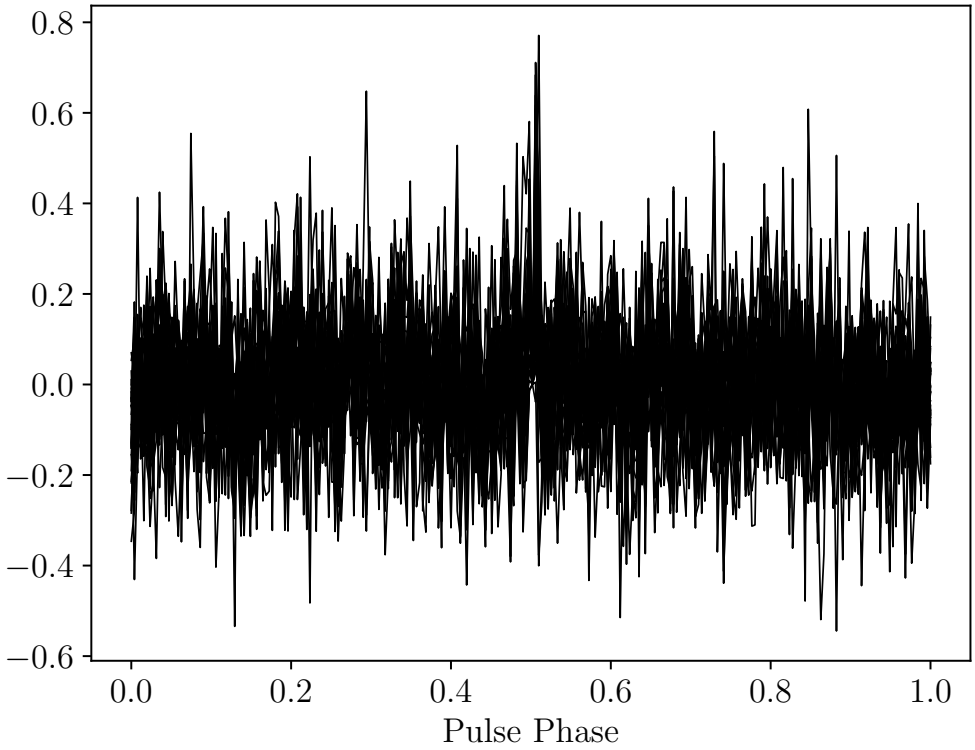


Figure 4.15: A difference plot, in 'maximum peak units', showing the raw flux (not Gaussian fit) offset of each observing day for PSR J1936+1805 with respect to the first timing epoch: 57924. Each day is superimposed on each other in order to compare baselines.

at a small spin-orbit misalignment angle (Farr et al., 2011) which supports evidence that suggests this system formed from a low-kick, symmetric supernova.

4.7 Summary

The pulsars showcased in this chapter represent the broad interests of the PALFA survey collaboration, as well as the wider pulsar community. This analysis has given rise to at least two potential double neutron star candidates (PSRs J1851+0010 and J1853+0008), adding to the growing tapestry of known compact binaries, with a third being a potential IMBP candidate — certainly all three binaries lie within both the expected $P_{\rm b} - P$ evolutionary region (Figure 4.8) and the expected $P - \dot{P}$ evolutionary region for their proposed binary categories. The mass of PSR J1851+0010's companion is potentially the lowest NS mass recorded to-date. This could have been due to a violent tidal stripping period or heavy mass-loss during its supernova. It is unlikely that the companion to PSR J1851+0010 is a WD considering the orbital parameters of the system being inconsistent with NS-WD evolution models. The survey has also unearthed an isolated pulsar whose rotation frequency and frequency derivative closely resemble that of a DNS binary (PSR J1936+1805) hinting that it is part of the population of so-called DRPs. Linking these types of disrupted binary systems with known compact binaries will hopefully give tighter constraints on the evolutionary channels that give rise to these exotic systems. Recently, the Canadian Hydrogen Intensity Mapping Experiment (CHIME; CHIME/Pulsar Collaboration et al., 2020) has begun observing PSRs J1851+0010 and J1853+0008 although, at the present time, it is unclear as to whether these pulsars are bright enough to be seen by CHIME. If there is a detectable pulsar signal in the data, it may give rise to proper motion measurements for these systems, although this is just speculation. If successful, this may help us to constrain the evolution of these systems.

Although it is clear that PSR J1936+2142 is a member of a short-period binary system, most likely comprising a NS and WD (possibly an IMBP); further analysis (from e.g. Green Bank Telescope) may aid in precisely determining the mass and proper motion parameters which will give key indicators as to its previous evolution. As previously mentioned, this pulsar's position on both the $P - \dot{P}$ diagram (Figure 4.7) and the $P_{\rm b} - P$ diagram (Figure 4.8) imply that it is closer in nature to the heavier binaries, such as the DNSs or the IMBPs, than it is to low-mass NS-WD binaries. Together with its relatively circular orbit and low Galactic scale height, an IMBP is the most likely system for this pulsar to reside in.

Chapter 5

Concluding remarks

The observed pulsar population shows a diverse range of spin and astrometric parameters. It is for this reason that current binary evolution models involving neutron stars are also rich and diverse but ultimately disconnected (e.g. Tauris et al., 2012; Tauris et al., 2017; van den Heuvel, 2007). Neutron stars exist in many different forms – some are isolated and slow, some are isolated and relatively fast; some exist in extremely circular binaries and others in very eccentric ones. Some have even been found with planets orbiting them (Wolszczan and Frail, 1992). The list of nuances to classification one could explore is endless, although some groupings amongst the population are more progressive than others (cf. e.g. Figures 3.11 and 3.14). What *is* clear is that there is still a gap in data which will be necessary to obtain in order to understand the true pathways of NS evolution. Adding to the pulsar "melting pot" through either long term high precision timing (Chapter 3) or targeted pulsar searches (Chapter 4) will eventually close those current gaps in understanding.

As radio pulsar timing becomes evermore precise, methods of treating data, such as removal of radio frequency interference and the accurate determination of pulse shape and flux, are inevitably becoming more sophisticated. To overcome the added computational cost that smart noise detection algorithms require, many researchers use rudimentary RFI excision algorithms and then 'clean' the rest of the data manually or some make use of server clusters for more expensive batch signal processing. The former is a problem as the accidental inclusion of specific regions of RFI can have a profound effect on pulse times-of-arrival. The latter is a problem due to the limited access many people have to those servers. In Chapter 2, the concept of 2D image recognition was applied to known 1D pulse profiles via a series of convolutional neural networks. These CNNs were designed with the previous issues in mind and subsequently tackled them via a three-pronged attack – heavily penalize strong results, compare results against more than one validator, and determine CNN architecture based on local machine memory capacities. This was done to good success although it is unclear at what level of precision the algorithms truly fail as robust tests have only been done on pulsars with $P > 20\,\mathrm{ms}$.

In Chapter 3, the compact binary evolution of DNS systems post-second-supernova was explored using the pulsar PSR J1829+2456 through an extensive pulsar timing campaign spanning over 17 years. This system is now among the few DNS binaries which

have precise component mass measurements. Component masses can give us a significant amount of information about the potential initial conditions of the stellar progenitors and their supernova explosions, the caveat being that only binary systems can hope to have measureable masses. The timing campaign has given another clue that the second-formed NS in this system probably evolved out of a low-kick symmetric supernova (much like e.g. PSRs J0737−3039B and J1756−2251; Kramer et al., 2006; Ferdman et al., 2014). This evolutionary subclass of DNS is becoming relatively more populous (than the other DNS subclass) as more component masses are measured. This may indicate a selection bias which allows mass measurements of certain types of orbit to be more feasible than with others. However, if the observed distribution of DNS types reflects a real trend, then measuring more DNS masses, orbital parameters and astrometric parameters will lead to a better understanding of these evolutionary mechanisms.

Probing the Universe for pulsar signals is of importance to many population and high-energy physics studies. Pulsars (namely NSs) are the conclusion of a series of violent reactions and unlikely interactions. Surveys that look towards the Galactic plane find many binary pulsars owing to the higher matter density at lower Galactic height. Chapter 4 discussed four new radio pulsar discoveries from the PALFA low-Galactic-plane survey (Cordes et al., 2006) in which three have been found to be in binary systems and the fourth pulsar having once been part of a binary system. The timing of PSRs J1851+0010 and J1853+0008 have shown them to have a relatively large total mass ($\sim 2.6\,M_\odot$) and a moderate eccentricity. This, along with their spin periods being in the tens of milliseconds and spin-down rates of $\sim 10^{-19}$, implies that these two systems are DNS systems, although component mass measurements for PSR J1853+0008 are still too uncertain to make definitive conclusions. If these are both DNS systems, the second-formed NS may also have formed like in the system containing PSR J1829+2456: with a low-kick, symmetric supernova from an iron core-collapse. The component masses measured for PSR J1851+0010 indicate that the companion could be the lowest-mass NS discovered to-date, although the mass still fits into an iron core-collapse scenario for the second supernova. This low mass is intriguing however as it puts tighter constraints on NS equations-of-state, given now the very wide mass ranges observed in DNS binaries, and puts constraints on scalarisation effects in the strong-field regime. The companion was most probably born with a mass even lower than this as the current mass is expected to have been developed after a period of mass accretion from the older NS. The birth mass associated with the companion to PSR J1851+0010 is hypothesised to be in the realms of $\approx 1\,M_\odot$ with an initial main sequence mass of $2-3\,M_\odot$.

The low eccentricity observed for PSR J1936+2142, along with its partially-recycled rotation period in the 10s of milliseconds, low upper transverse velocity estimate and low Galactic scale height, imply that this system is probably an IMBP: a population of NS-WD binary systems such that $M_{\rm WD} \gtrsim 0.4\,M_\odot$ in which relatively few component masses have been measured. Unfortunately, the 50+ year mission at Arecibo has come to an end and, with it, the life of many of the PALFA discovered pulsars, which are potentially not bright enough at any other site for follow-up timing. The CHIME project (CHIME/Pulsar Collaboration et al., 2020) has so far attempted to observe many PALFA-discovered pulsars,

including PSR J1851+0010 and PSR J1853+0008, however there are currently no published results for these projects.

The field of radio astronomy is by no means slowing down – if anything, the radio Universe becomes evermore alive the more precise our instrumentation gets. Newly discovered, previously inexplicable, phenomena such as fast radio bursts (FRBs; Keane et al., 2012), some of which have been observed to repeat (e.g. FRB121102; Cruces et al., 2020), and other transient species only add to the mysteries of the Universe. Even more recently, a FRB was claimed as being detected from within our own Galaxy which has led to the idea that at least some FRBs are magnetars (Bochenek et al., 2020). In any case, the proposed evolution mechanisms governing the currently observed pulsar population are poorly understood to the level of detail that these tests now demand. As a community, we have a base understanding of binary evolution that correlates reasonably well with the observed population of stellar remnant binaries, X-ray binaries, and high-mass main sequence binaries, however as these evolution categories become narrower in scope, so too do the necessary minimum constraints on measureable binary parameters. It is the hope that this work has contributed to tightening these contraints to some degree, even if all of the systems discussed lie within currently accepted evolutionary pathways. After all, as we climb higher, so too does the base platform beneath our feet.

Appendix A

Derivations and methods

A.1 Braking index

Assume that the change in pulse frequency can be expressed as a power law in the frequency:

$$\dot{\nu} = k\nu^n$$

where n is defined as the braking index and k is a constant. Taking the second derivative of the frequency, one finds:

$$\begin{aligned}
\ddot{\nu} &= kn\nu^{n-1}\dot{\nu} \\
&= n\left(\frac{k\nu^n}{\nu}\right)\dot{\nu} \\
&= n\left(\frac{\dot{\nu}}{\nu}\right)\dot{\nu} \\
&= \frac{n\dot{\nu}^2}{\nu}
\end{aligned}$$

In terms of n:

$$n = \frac{\nu\ddot{\nu}}{\dot{\nu}^2} \tag{A.1}$$

Since $\nu = P^{-1}$, $\dot{\nu} = -P^{-2}\dot{P}$ and $\ddot{\nu} = -P^{-2}\ddot{P} + 2P^{-3}\dot{P}^2$, combining these with Equation A.1, one arrives at the braking index in terms of the period:

$$\begin{aligned}
n &= \left(\frac{1}{P}\right)\left(\frac{2\dot{P}^2}{P^3} - \frac{\ddot{P}}{P^2}\right)\left(-\frac{\dot{P}}{P^2}\right)^{-2} \\
&= 2 - \frac{P\ddot{P}}{\dot{P}^2}
\end{aligned} \tag{A.2}$$

A.2 χ^2 formulation

Using the formulation by Lancaster (Lancaster, 1969), first consider the binomial distrubition which can be well approximated by a normal distribution. Specifically, for m observed successes occurring with probability p over N trials, the asymptotic case in normality for the binomial distribution, χ, can be given by:

$$\chi = \frac{m - Np}{\sqrt{Npq}} \tag{A.3}$$

where q is the probability of failure ($q \equiv 1 - p$). Squaring both sides gives:

$$\chi^2 = \frac{(m - Np)^2}{Npq} \tag{A.4}$$

Converting this to partial fractions, and using the identities $N = Np + N(1-p) = Np + Nq$ and $N = m + (N - m)$, one arrives at:

$$\chi^2 = \frac{(m - Np)^2}{Np} + \frac{(N - m - Nq)^2}{Nq} \tag{A.5}$$

Extending this to a distribution of n types of observation (as opposed to just one), the expression becomes:

$$\chi^2 = \sum_{i=1}^{n} \frac{(m_i - Np_i)^2}{Np_i} \tag{A.6}$$

where m_i is the observation value and Np_i is the theoretical value for observation i. The numerators in the above expression are identical to the model residuals squared. In pulsar astronomy, the *reduced* χ^2 over N TOAs is given by summing the pre-fit residuals, R_i, divided by the error in those residuals, σ_i, assuming all observations are independent. This is further divided by the degrees of freedom in model, Γ:

$$\chi^2_{\mathrm{r}} = \frac{1}{\Gamma} \sum_{i=1}^{N_{\mathrm{TOA}}} \left(\frac{R_i}{\sigma_i} \right)^2 \tag{A.7}$$

A.3 Activation function derivatives

Sigmoid

$$\sigma(x) = (1 + e^{-x})^{-1}$$

Letting $u = 1 + e^{-x}$, one then has:

$$\begin{aligned}\sigma'(x) &= -u^{-2}u' \\ &= e^{-x}(1 + e^{-x})^{-2} \\ &= e^{-x}\sigma^2(x)\end{aligned}$$

Noticing that $1 + e^{-x} = 1/\sigma(x)$, the above can be transformed into:

$$\begin{aligned}\sigma'(x) &= \left(\frac{1 - \sigma(x)}{\sigma(x)}\right)\sigma^2(x) \\ &= \sigma(x) - \sigma^2(x) \\ &= \sigma(x)(1 - \sigma(x))\end{aligned}$$

Hyperbolic tangent

via quotient rule.

$$\begin{aligned}\tanh(x) &= 2(1 + e^{-2x})^{-1} - 1 \\ &= \frac{2 - (1 + e^{-2x})}{1 + e^{-2x}} \\ &= \frac{1 - e^{-2x}}{1 + e^{-2x}} = \frac{e^{2x} - 1}{e^{2x} + 1} \\ &= \frac{e^x - e^{-x}}{e^x + e^{-x}} \\ \therefore [\tanh(x)]' &= \frac{(e^x + e^{-x})(e^x + e^{-x}) - (e^x - e^{-x})(e^x - e^{-x})}{(e^x + e^{-x})^2} \\ &= 1 - \frac{(e^x - e^{-x})^2}{(e^x + e^{-x})^2} \\ &= 1 - \tanh^2(x)\end{aligned}$$

Rectified linear unit

From the definition of $\mathrm{ReLU}(x)$:

$$\mathrm{ReLU}(x) = \max(0, x)$$

the range can be split into two regions of $x < 0$ and $x > 0$ with one pole at $= 0$. The former region is not dependent on x so has a derivative of 0. In the latter region, $\mathrm{ReLU}(x) = x$

so the gradient is 1. At the pole it is undefined.

Softmax

From Bendersky's derivation[1]:

$$\mathrm{s}(\mathbf{x}) = \frac{e^{\mathbf{x}}}{\sum_{k=1}^{N} e^{x_k}}$$

Let $\Sigma = \sum_{k=1}^{N} e^{x_k}$. Then, for $i = j$:

$$\begin{aligned}
\frac{\partial \mathrm{s}_i}{\partial q_j} &= \frac{\partial \frac{e^{x_i}}{\Sigma}}{\partial x_j} \\
&= \frac{e^{x_i}\Sigma - e^{x_j}e^{x_i}}{\Sigma^2} \\
&= \frac{e^{x_i}}{\Sigma}\frac{\Sigma - e^{x_j}}{\Sigma} \\
&= \mathrm{s}_i(1 - \mathrm{s}_j)
\end{aligned}$$

Similarly for $i \neq j$:

$$\begin{aligned}
\frac{\partial \frac{e^{x_i}}{\Sigma}}{\partial x_j} &= \frac{0 - e^{x_j}e^{x_i}}{\Sigma^2} \\
&= \frac{-e^{x_j}}{\Sigma}\frac{e^{x_i}}{\Sigma} \\
&= -\mathrm{s}_j\mathrm{s}_i
\end{aligned}$$

Combining these two cases, one obtains:

$$\frac{\partial \mathrm{s}_i}{\partial q_j} = \mathrm{s}_i(\delta_{ij} - \mathrm{s}_j)$$

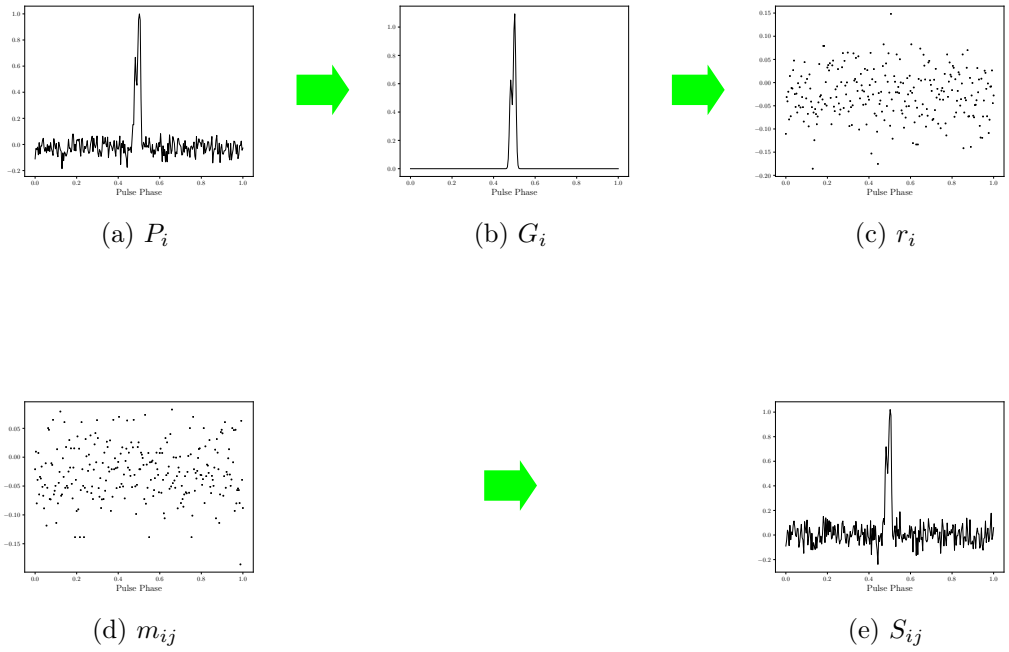
where δ is the Kronecker delta function.

Swish

via product rule.

$$\begin{aligned}
\varsigma(x) &= x\sigma(bx) \\
\therefore \varsigma'(x) &= \sigma(bx) + x\sigma(bx)\left(1 - \sigma(bx)\right) \\
&= \varsigma(x) - \left(x\sigma^2(bx) - \sigma(bx)\right) \\
&= \varsigma(x) - \left(\varsigma(x)\sigma(bx) - \sigma(bx)\right) \\
&= \varsigma(x) + \sigma(bx)\left(1 - \varsigma(x)\right)
\end{aligned}$$

[1] https://eli.thegreenplace.net/2016/the-softmax-function-and-its-derivative/



(a) P_i (b) G_i (c) r_i

(d) m_{ij} (e) S_{ij}

Figure A.1: Flowchart outlining the bootstrap algorithm.

A.4 Bootstrap algorithm

The bootstrap algorithm for finding measurement uncertainties (see Section 4.6) is given as follows (with examples):

Integrated pulse profiles, P_i, of b number of data points are made for each date of observation, i. These are fit to a Gaussian function, G_i. Profile residuals, r_i, are calculated by subtracting the Gaussian fit from the observed integrated profile. From these residuals, b values are chosen at random with replacement using a uniform distribution to create a new set of residuals, m_{ij}, where j is the iteration number. These model residuals are then subtracted from P_i in order to obtain a statistically created profile, S_{ij}. In the case of this thesis, the on-pulse widths at 10% and 50% of the maximum is calculated from this model profile. The random selection part (and subsequent steps) is done n times. An example of this procedure is given in Figure A.1.

The widths from n iterations are then plotted on a histogram and fit to a single Gaussian (see Figure A.2). The reported uncertainty in the width measurement of P_i is then taken to be the 2σ (95%) confidence interval of the histogram Gaussian fit.

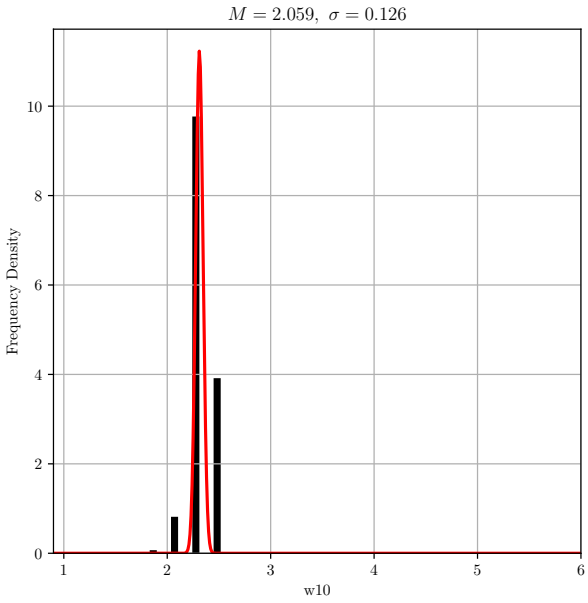


Figure A.2: Example histogram of 10% pulse widths for PSR J1936+1805. The Gaussian fit is given in red.

Appendix B

Supplementary information

B.1 Standard profile breakdown

The general theory behind Gaussian profile summation can be found in Foster et al. (1991). Gaussians take the general form:

$$g(\mathbf{x}) = ae^{\frac{(\mathbf{x}-b)^2}{2c^2}} \tag{B.1}$$

for vector, $\mathbf{x}$, with a, b and c corresponding to the shape parameters in the following tables. All values are given in arbitrary flux on the y-axis and in phase bins on the x-axis, except in the instances of the fold mode templates for PSRs J1851+0010 and J1853+0008, where the `pav` output displays the templates in terms of fractional pulse phase. All profiles have been centered for the plots but their original center fitting value is preserved in the tabulation.

PSR J1829+2456

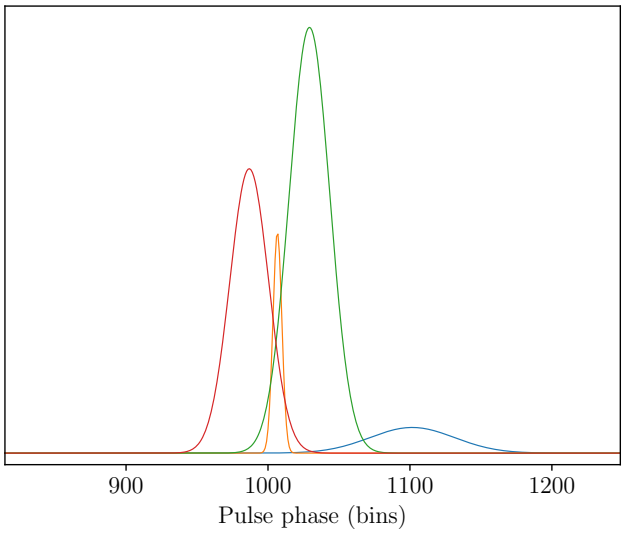


Figure B.1: Individual Gaussian curves for the L–band standard profile of PSR J1829+2456 fit from data spanning dates 57950 − 58948.

Table B.1: Parameters for the individual Gaussians comprising the L–band standard profile of PSR J1829+2456.

$g(\mathbf{x})$	a	b	c
1	81608.5	1101.6	29.779
2	73995.2	1006.6	3.140
3	652525	1029.2	14.356
4	414676	986.8	13.658

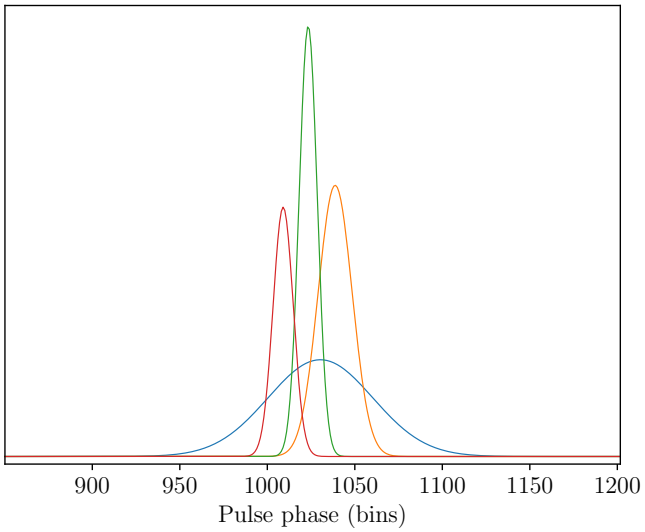


Figure B.2: Individual Gaussian curves for the 430 MHz standard profile of PSR J1829+2456 fit from data spanning dates 57950 – 58948.

Table B.2: Parameters for the individual Gaussians comprising the standard profile of PSR J1829+2456.

$g(\mathbf{x})$	a	b	c
1	0.582	1030.3	29.94
2	0.528	1038.8	9.704
3	0.446	1023.4	5.164
4	0.287	1009.1	5.736

PSR J1851+0010

The raw integrated profile (uncentred):

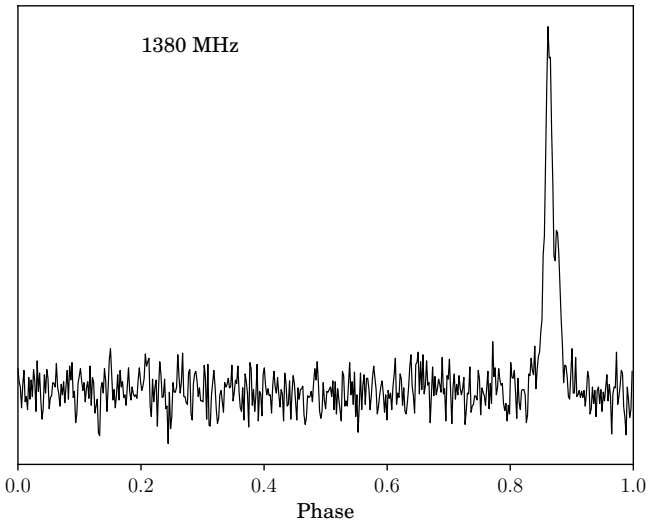


Figure B.3: The L-band raw integrated profile for PSR J1851+0010, averaged from data spanning dates 57330 – 58680.

The search mode template for PSR J1851+0010 has a three-Gaussian fit where maybe one or two Gaussians would have sufficed. As such, the profile is not shown as it is practically identical to Figure 4.4, panel 1. Also included here is the fold mode template used, created from all available fold mode observations using `psradd` in PSRchive.

Table B.3: Parameters for the individual Gaussians comprising the standard profile of PSR J1851+0010.

$g(\mathbf{x})$	a	b	c
1	4667.5	450.8	3.589
2	4639.0	450.8	3.587
3	4777.9	450.8	3.589

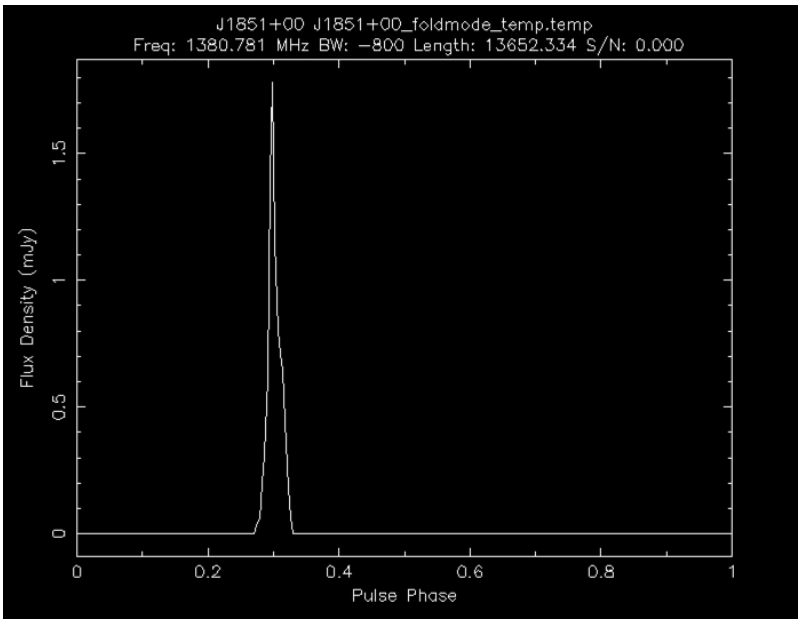


Figure B.4: The L-band fold mode template for PSR J1851+0010, summed together from data spanning dates 57330 − 58680 and fit using routines in PSRchive.

PSR J1853+0008

The raw integrated profile (uncentred):

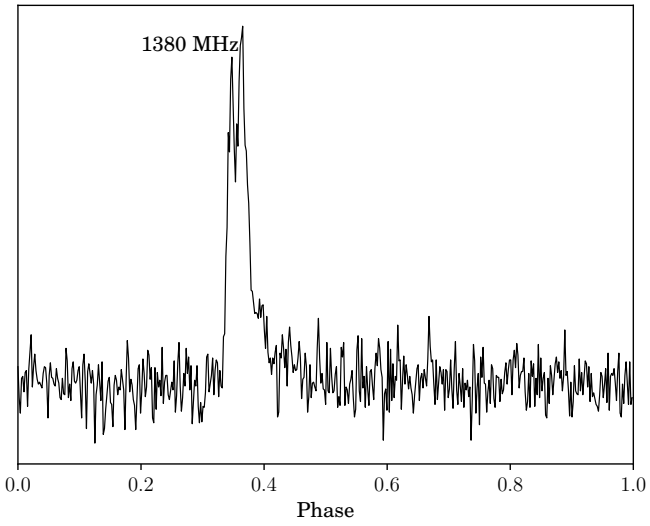


Figure B.5: The L-band raw integrated profile for PSR J1853+0008, averaged from 2 hours of data from MJD 57303.

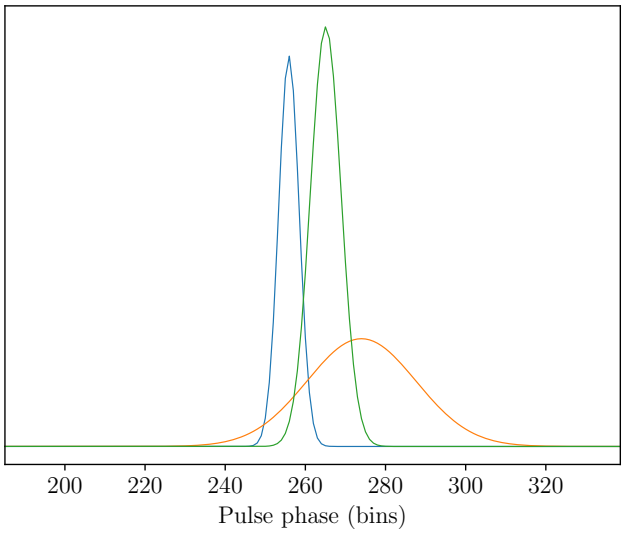


Figure B.6: Individual Gaussian curves for the L-band search mode template of PSR J1853+0008, fit from data taken on 57303.

Table B.4: Parameters for the individual Gaussians comprising the standard profile of PSR J1853+0008.

$g(\mathbf{x})$	a	b	c
1	171394	176.9	2.570
2	250356	195.0	13.61
3	274232	186.0	3.827

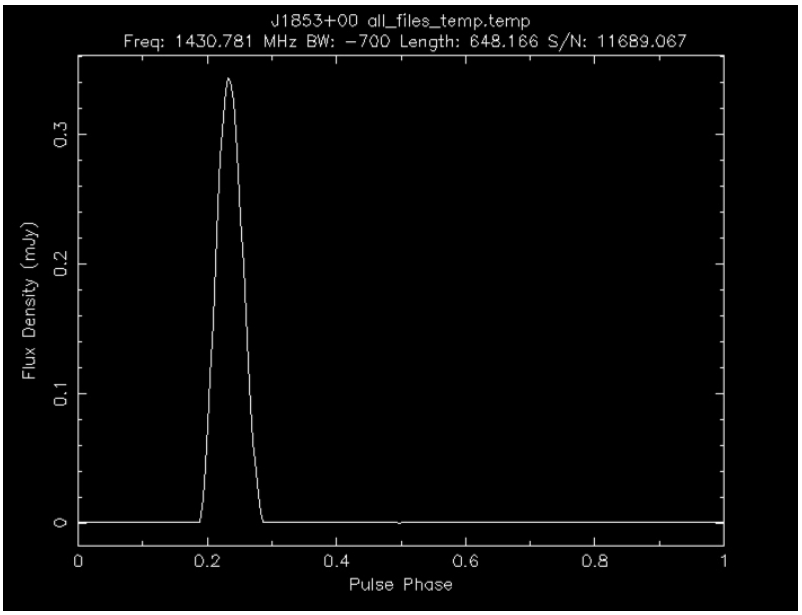


Figure B.7: The L-band fold mode template for PSR J1853+0008, summed together from data spanning dates 57330 – 58680 and fit using routines in PSRchive.

PSR J1936+1805

The raw integrated profile (uncentred):

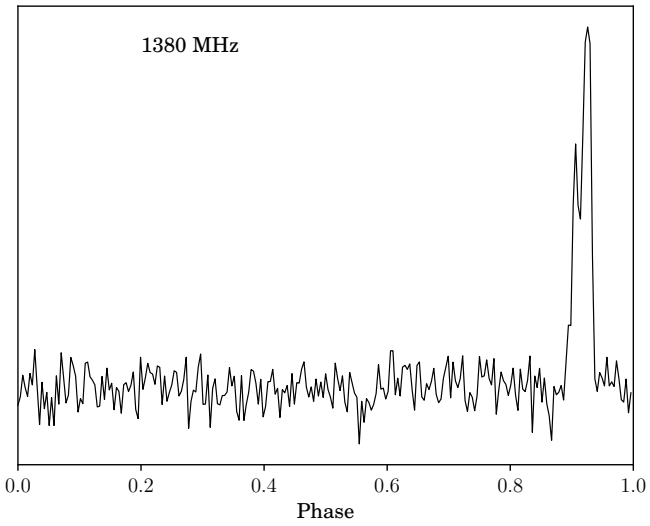


Figure B.8: The L-band raw integrated profile for PSR J1936+1805, averaged from 1 hour of data from MJD 57924.

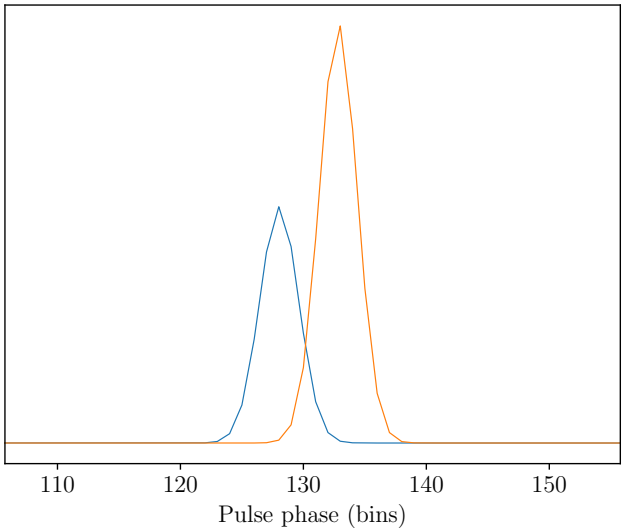


Figure B.9: Individual Gaussian curves for the L-band template of PSR J1936+1805, fit from search mode data taken on 57924.

Table B.5: Parameters for the individual Gaussians comprising the standard profile of PSR J1936+1805.

$g(\mathbf{x})$	a	b	c
1	440325	232.0	1.587
2	754052	236.8	1.531

PSR J1936+2142

The raw integrated profile (uncentred):

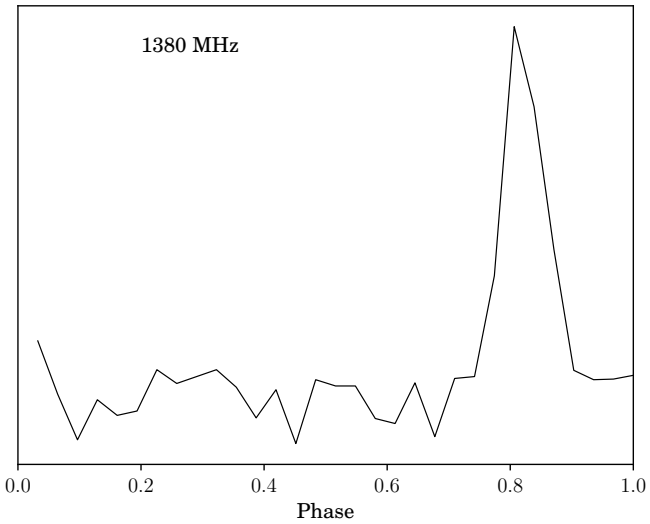


Figure B.10: The L-band raw integrated profile for PSR J1936+2142, averaged from 1 hour of data from MJD 57679.

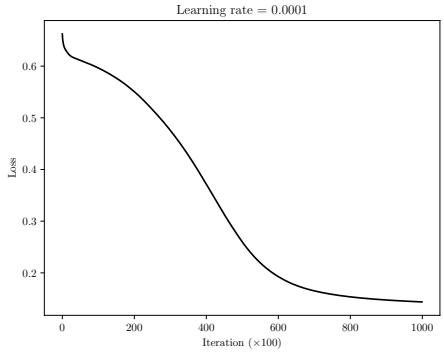
The template file for PSR J1936+2142 was fit to one Gaussian; as such, a repeat of Figure 4.4, panel 4, is not necessary. Detailed below are the Gaussian parameters that give that panel.

Table B.6: Parameters for the individual Gaussians comprising the standard profile of PSR J1936+2142, which is based on the integrated profile from an observation taken on date 57679.

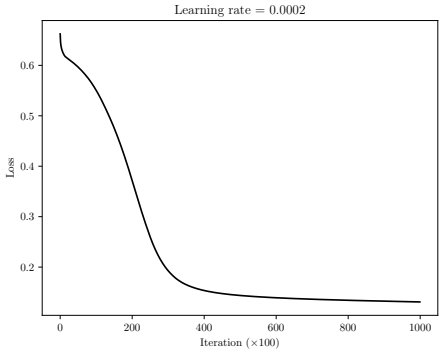
$g(\mathbf{x})$	a	b	c
1	153777	51.37	2.111

B.2 Loss profiles for NN optimisation

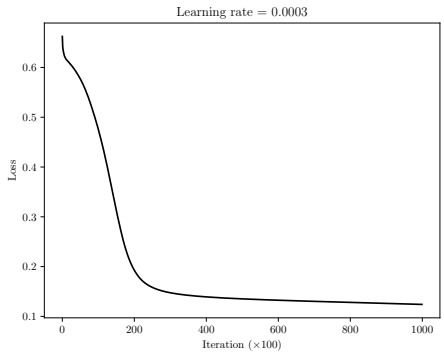
The following section documents the post-training loss profiles that lead to the loss values as reported in Table 2.2. The plots are organised by the specific CNN activation architecture with each subplot corresponding to a specific learning rate. All training was conducted over $100,000$ iterations with a threshold of 0.7. The 'bumpy' nature of the trailing end of some profiles gives the learning ability limit (the iteration at which the loss displays chaotic behavior) of the setup.



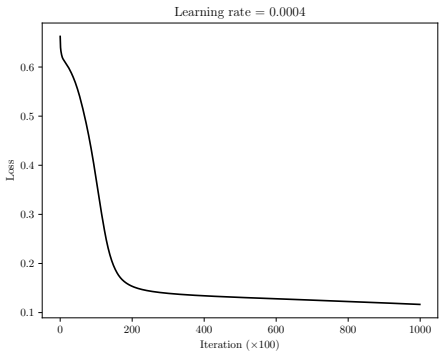
(a)



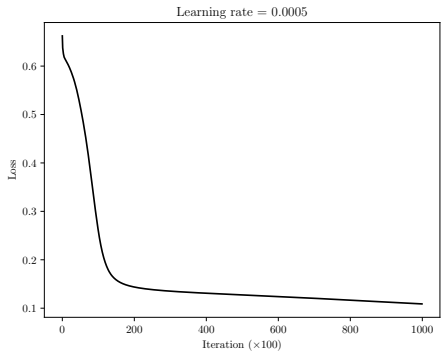
(b)



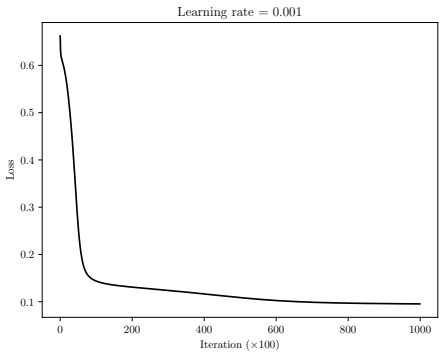
(c)



(d)

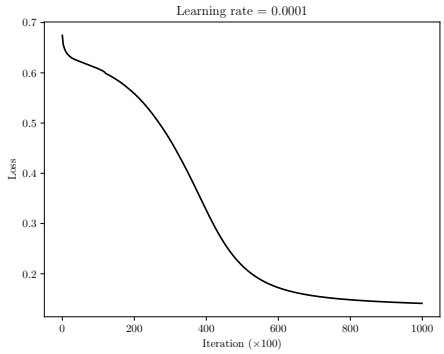


(e)

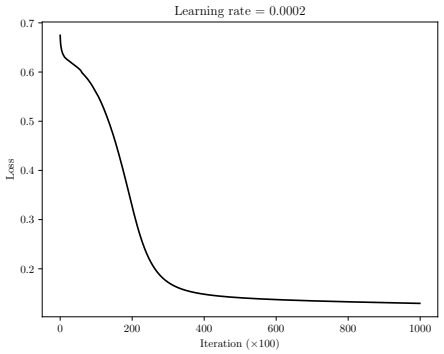


(f)

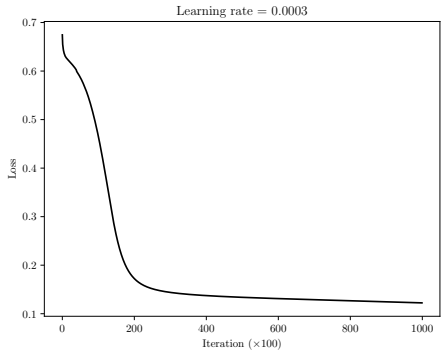
Figure B.11: Loss profiles for ReLU$(x), \sigma(x)$ CNN architecture.



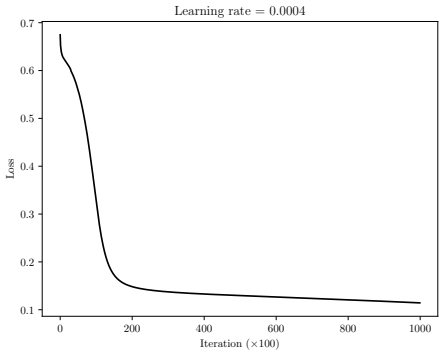
(a)



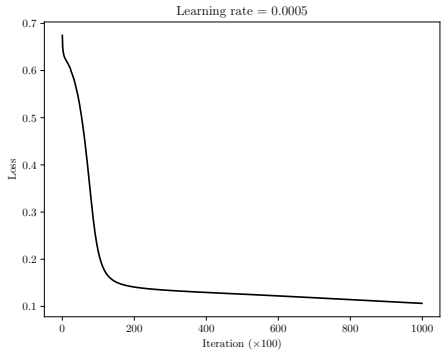
(b)



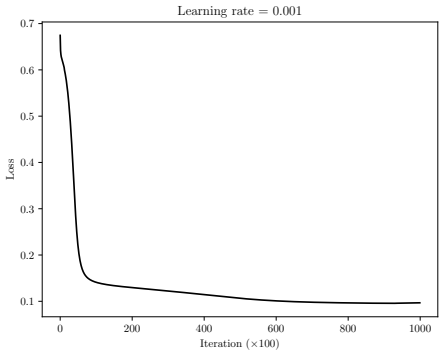
(c)



(d)

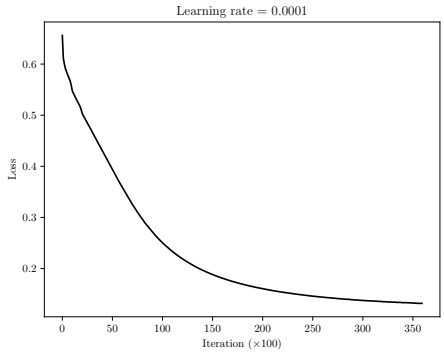


(e)

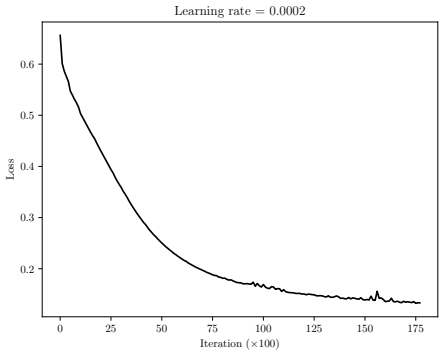


(f)

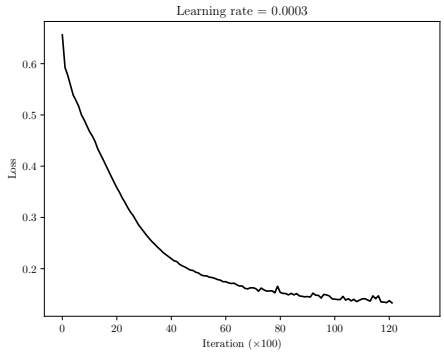
Figure B.12: Loss profiles for $\varsigma(x), \sigma(x)$ CNN architecture.



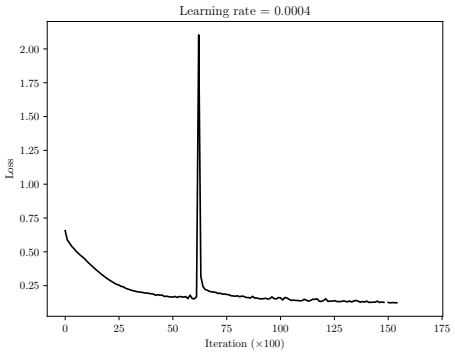
(a)



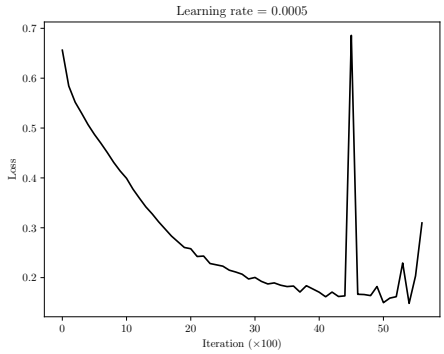
(b)



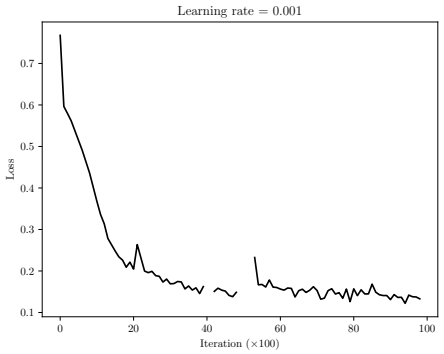
(c)



(d)

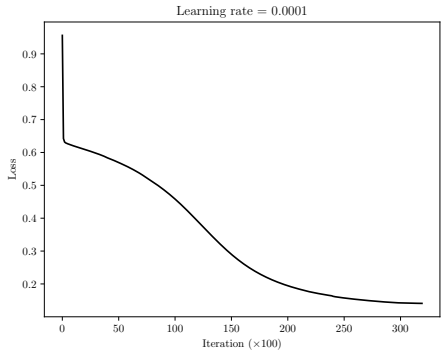


(e)

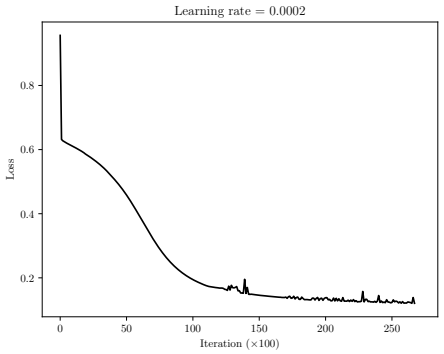


(f)

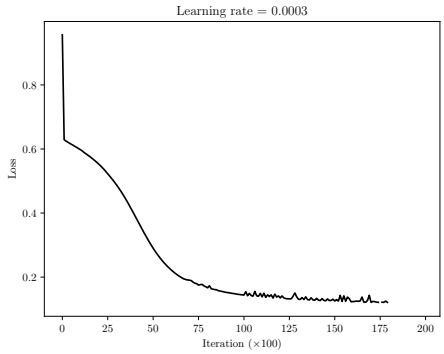
Figure B.13: Loss profiles for $\mathrm{ReLU}(x), \tanh(x)$ CNN architecture.



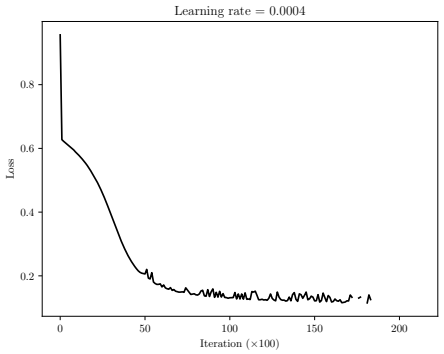
(a)



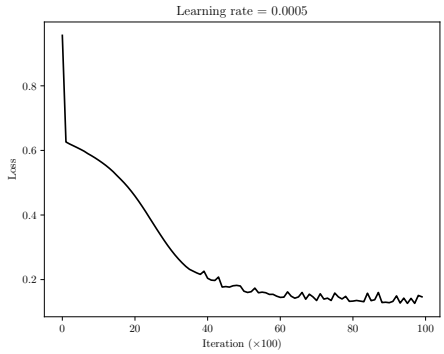
(b)



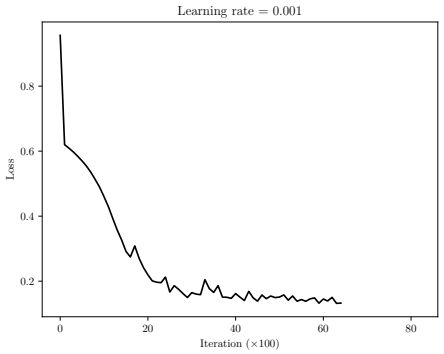
(c)



(d)

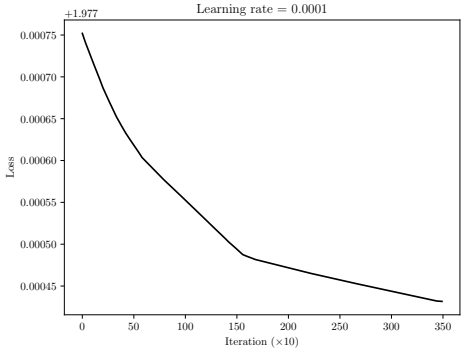


(e)

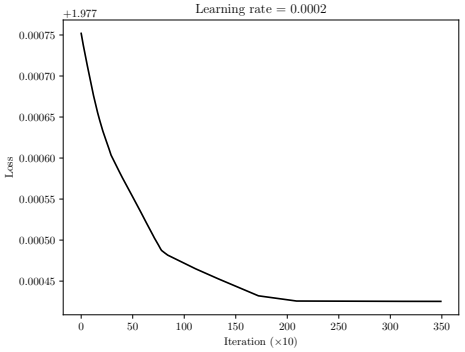


(f)

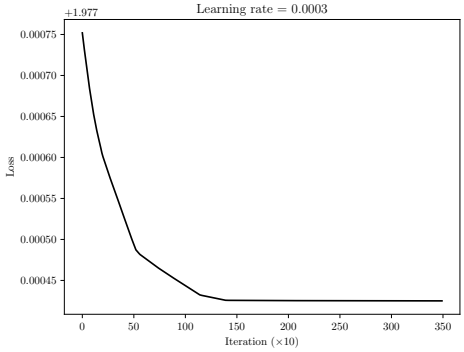
Figure B.14: Loss profiles for $\varsigma(x), \tanh(x)$ CNN architecture.



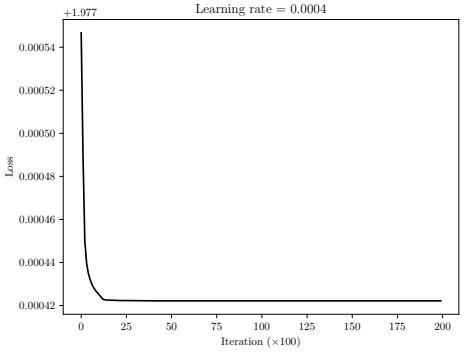
(a)



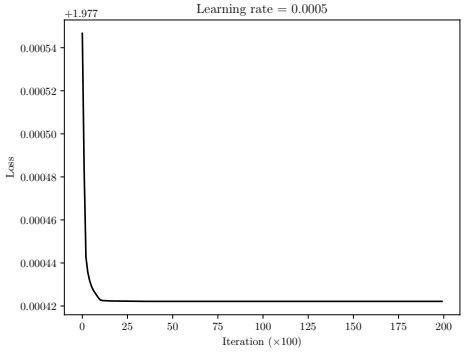
(b)



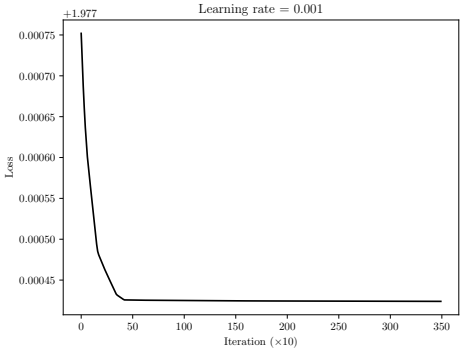
(c)



(d)

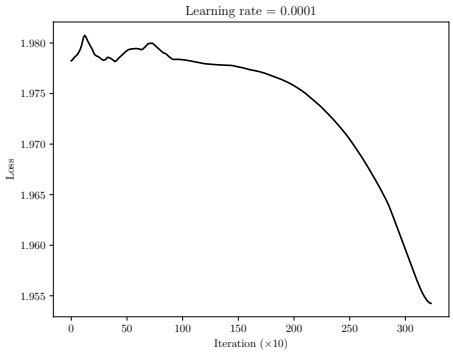


(e)

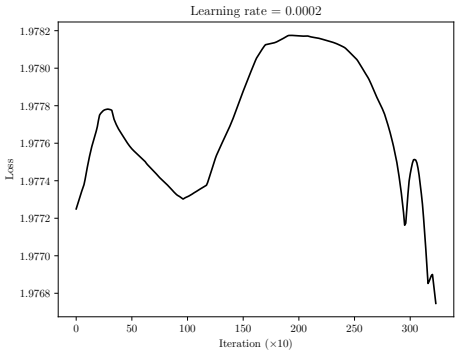


(f)

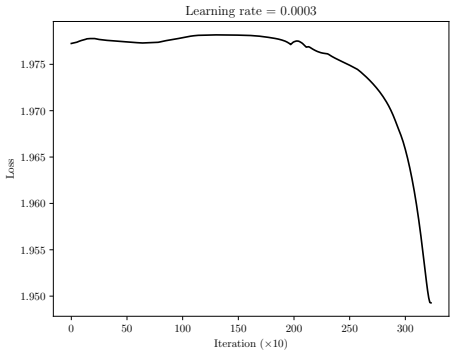
Figure B.15: Loss profiles for ReLU(x), s$(\mathbf{x})$ CNN architecture.



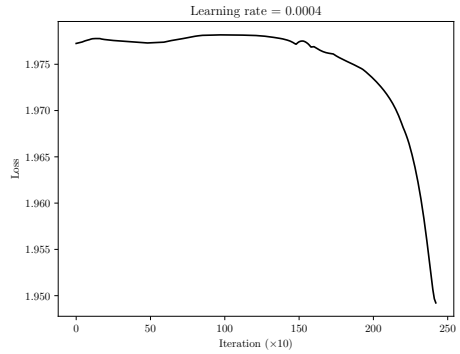
(a)



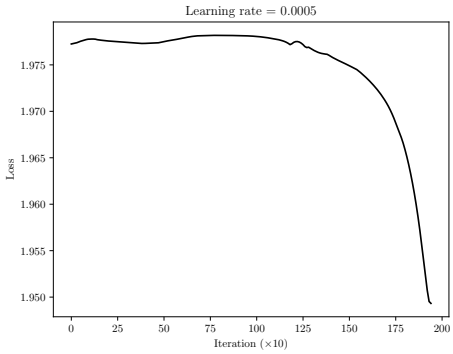
(b)



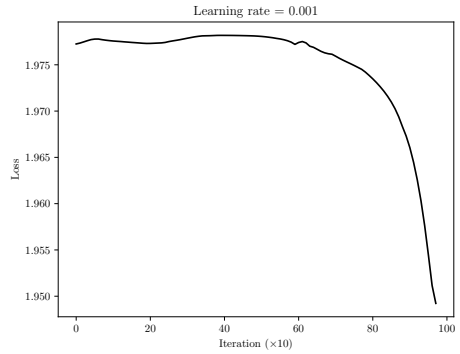
(c)



(d)



(e)



(f)

Figure B.16: Loss profiles for $\varsigma(x), \mathrm{s}(\mathbf{x})$ CNN architecture.

Appendix C

PSRVoid

The software built for the purposes of this work is called PSRVoid[1] (Haniewicz, 2020) and is responsible for the RFI exicision (via all the methods discussed in Chapter 2), profile flux calibration and time-of-arrival calculations. It is written in Python 3.8 but has a lot of backward compatibility with Python 2 outputs. PSRVoid's main I/O infrastructure is based on the PSRFITS processing package PyPulse (Lam, 2017) however the data manipulation methods employed in PSRVoid deviate from PyPulse's current implementations. PSRVoid is a fast, local alternative to processes that would usually be done on more specialist servers.

Flux calibration is administered on fold mode data via the `cal.py` program in the following syntax: `cal.py [file] [continuum] [continuum-directory] [verbose]`. Standard setup for flux calibration using `cal.py` is via Format 2 (see Chapter 2) with a telescope gain of $11.0\,\mathrm{K\,Jy^{-1}}$ although these values can be changed by the user. Interpolation for the least-squares fit is conducted via a 1D cubic spline in SciPy's `interpolate` package. The output is an ASCII file with five column headers: Frequency (in MHz), A^2, B^2, CR, CI; the latter four values being given in Jy/count. This ASCII file can then be loaded on the original file by PSRVoid ready for RFI excision.

RFI excision is conducted via the `rfi.py` program with neural network training being done via `nn.py`. The former has the syntax: `rfi.py [file] [template] [method] [cal]`; and the latter: `nn.py [training-file] [validation-file]`, although in both cases, many options (such as which curves to attempt to fit to histograms) are customisable by the user within the code. The neural network training outputs a series of NumPy arrays each representing a layer of the neural network which can be loaded into a feed-forward routine in `rfi.py` in order to determine which frequency channels to omit. `rfi.py` can also be used for more traditional auto-RFI excision techniques such as Gaussian noise distribution σ-clipping. The output of `rfi.py` is an ASCII file containing a list of subintegrations and frequencies (in MHz) to omit when calculating TOAs.

PSRVoid has many tools for the average radio astronomer. These include: a data plotter which is capable of not only plotting profiles in phase space, but can also plot off-pulse RMS in time-frequency space; numerous data miners for both the ATNF Pulsar Database (Manchester et al., 2005) and other useful astrophysical websites; and routines

[1] https://github.com/HenrykHaniewicz/PSRVoid – not pip installable

to estimate kinematic effects in pulsar timing residuals. PSRVoid also contains a Perl script to convert right ascension and declination into galactic coordinates.

The aim of PSRVoid is to show the breadth of what a pulsar suite can potentially do. One can use it as a complete software package to perform flux calibration, RFI excision and timing or as a set of functions and routines to use in further work (this was indeed the *initial* goal of PSRVoid). The latter prospect allows PSRVoid to act as an arbitrary data miner – it can even be used to mine cryptocurrency with some small tweaks (although this is not its intended functionality), and the routines can certainly be used to mine resources from other pulsar websites that are not the ATNF (such as https://www3.mpifr-bonn.mpg.de/staff/pfreire/NS_masses.html).

The future of PSRVoid is uncertain although one idea I have had is as follows: PSRVoid could be run in the background when the computer is booted, using extremely low memory, with the program constantly searching through given directories for any files it has not yet analysed. Once a FITS file is downloaded to one of these directories, it would eventually be seen by PSRVoid which would then read the FITS file, sort it into a correct pulsar folder (or create a new one), perform relevant data analysis (calibration, RFI excision, etc.) and, if it determines there are enough TOAs, attempt to create / update the timing model. Then, all the user would have to do would be to see the output timing models after a few hours and possibly add tweaks if necessary. This would leave more time for the user to tackle the more important astrophysical implications of the data instead of spending time devising timing models. This is ambitious but is already in its alpha stage of development.

Bibliography

Abbott, B. P. et al. (2017). *The Astrophysical Journal Letters* 848 (2), p. L12.

Abramowitz, M. and Stegun, I. A. (1972).

Allen, B. et al. (2013). *The Astrophysical Journal* 773 (2), p. 91.

Alpar, M. A. et al. (1982). *Nature* 300 (5894), pp. 728–730.

Anderson, D. P., Christensen, C., and Allen, B. (2006). *Proceedings of the 2006 ACM/IEEE Conference on Supercomputing*. Tampa, FL: Association for Computing Machinery, pp. 126–136.

Antoniadis, J. et al. (2016). *arXiv: High Energy Astrophysical Phenomena*.

Archibald, A. M. et al. (2018). *Nature* 559 (7712), pp. 73–76.

Armstrong, E. S. (1921). *Proceedings of the Institute of Radio Engineers* 9, pp. 3–11.

Arzoumanian, Z. et al. (2018). *The Astrophysical Journal Supplement Series* 235 (2), p. 37.

Baade, W. and Zwicky, F. (1934a). *Proceedings of the National Academy of Sciences* 20 (5), pp. 259–263.

— (1934b). *Proceedings of the National Academy of Sciences* 20 (5), pp. 254–259.

— (1934c). *Physical Review* 46, pp. 76–77.

Baars, J. W. M. et al. (1977). *Astronomy and Astrophysics* 500, pp. 135–142.

Baevski, A. et al. (2020). *arXiv e-prints*.

Bahcall, J. N. and Ostriker, J. P. (1997). Princeton University Press.

Bailes, M. (1989). *The Astrophysical Journal* 342, p. 917.

Belczynski, K., Kalogera, V., and Bulik, T. (2002). *The Astrophysical Journal* 572 (1), pp. 407–431.

Bell, J. F. et al. (1995). *The Astrophysical Journal Letters* 447, p. L117.

Bhattacharya, D. and van den Heuvel, E. P. J. (1991). *Physics Reports* 203 (1-2), pp. 1–124.

Bhattacharyya, B. et al. (2016). *The Astrophysical Journal* 817 (2), p. 130.

Bisnovatyi-Kogan, G. S. and Komberg, B. V. (1974). *Soviet Astronomy* 18, p. 217.

Blandford, R. D. and Romani, R. W. (1988). *Monthly Notices of the Royal Astronomical Society* 234, 57P–60.

Blandford, R. and Teukolsky, S. A. (1976). *The Astrophysical Journal* 205, pp. 580–591.

Bochenek, C. D. et al. (2020). *Publications of the Astronomical Society of the Pacific* 132 (1009).

Bolton, S. J. (2010). *Proceedings of the International Astronomical Union* 6 (S269), pp. 92–100.

Boyles, J. et al. (2013). *The Astrophysical Journal* 763 (2), p. 80.

Brinkman, C. et al. (2017). *Monthly Notices of the Royal Astronomical Society* 474 (2), pp. 2012–2027.

Burgay, M. et al. (2003). *Nature* 426 (6966), pp. 531–533.

Caballero, R. N. et al. (2018). *Monthly Notices of the Royal Astronomical Society* 481 (4), pp. 5501–5516.

Cameron, A. D. et al. (2018). *Monthly Notices of the Royal Astronomical Society: Letters* 475 (1), pp. L57–L61.

Camilo, F., Nice, D. J., and Taylor, J. H. (1993). *Astrophysical Journal Letters* 412, p. L37.

Camilo, F., Thorsett, S. E., and Kulkarni, S. R. (1994). *The Astrophysical Journal Letters* 421, p. L15.

Camilo, F. et al. (2001). *The Astrophysical Journal* 548 (2), pp. L187–L191.

Caruana, R., Lawrence, S., and Giles, C. (2000). Vol. 13, pp. 402–408.

Chadwick, J. (1932). *Proc. R. Soc. Lond. A*. 136.

Champion, D. J. et al. (2004). *Monthly Notices of the Royal Astronomical Society* 350, pp. 61–65.

Champion, D. J. et al. (2005). *Monthly Notices of the Royal Astronomical Society* 363, pp. 929–936.

Champion, D. J. et al. (2010). *The Astrophysical Journal* 720 (2), pp. L201–L205.

Chandrasekhar, S. (1947). *Bulletin of the American Mathematical Society* 53 (7), pp. 641–711.

Chaurasia, H. K. and Bailes, M. (2005). *The Astrohysical Journal* 632 (2), pp. 1054–1059.

Chauvenet, W. (1863). 5th. Vol. 5. 2, pp. 474–566.

CHIME/Pulsar Collaboration et al. (2020). *arXiv e-prints*, arXiv:2008.05681.

Coles, W. et al. (2011). *Monthly Notices of the Royal Astronomical Society* 418 (1), pp. 561–570.

Contopoulos, I. and Spitkovsky, A. (2006). *The Astrophysical Journal* 643, pp. 1139–1145.

Cordes, J. M. (1978). *The Astrophysical Journal* 222, pp. 1006–1011.

— (2002). *Single-Dish Radio Astronomy: Techniques and Applications*. Ed. by S. Stanimirovic et al. Vol. 278. Astronomical Society of the Pacific Conference Series, pp. 227–250.

Cordes, J. M. and Lazio, T. J. W. (2002), astro–ph/0207156.

— (2003).

Cordes, J. M. et al. (2006). *The Astrophysical Journal* 637 (1), pp. 446–455.

Corongiu, A. et al. (2007). *Astronomy and Astrophysics* 462 (2), pp. 703–709.

Craft, H. D. and Comella, J. M. (1968). *Nature* 220 (5168), pp. 676–678.

Cromartie, H. T. et al. (2019). *Nature Astronomy* 4 (1), pp. 72–76.

Cruces, M. et al. (2020). *Monthly Notices of the Royal Astronomical Society* 500 (1), pp. 448–463.

Curry, H. B. (1944). *Quarterly of Applied Mathematics* 2 (3), pp. 258–261.

Damour, T. and Deruelle, N. (1985). *Ann. Inst. Henri Poincaré Phys. Théor* 43 (1), pp. 107–132.

— (1986). *Ann. Inst. Henri Poincaré Phys. Théor* 44 (3), pp. 263–292.

Damour, T. and Esposito-Farèse, G. (1992). *Classical and Quantum Gravity* 9 (9), pp. 2093–2176.

— (1993). *Phys. Rev. Lett.* 70 (15), pp. 2220–2223.

— (1996). *Physical Review D* 54 (2), pp. 1474–1491.

de Araujo, J. C. N., Coelho, J. G., and Costa, C. A. (2016). *Journal of Cosmology and Astroparticle Physics* 2016 (07), pp. 023–023.

Deneva, J. S. et al. (2013). *The Astrophysical Journal* 775 (1), p. 51.

Devine, T. R., Goseva-Popstojanova, K., and McLaughlin, M. (2016). *Monthly Notices of the Royal Astronomical Society* 459 (2), pp. 1519–1532.

Dewey, R. J. (1992). *X-Ray Binaries and Recycled Pulsars* 377, pp. 453–463.

Dewi, J. D. M., Podsiadlowski, P., and Sena, A. (2006). *Monthly Notices of the Royal Astronomical Society* 368 (4), pp. 1742–1748.

Dyks, J. (2017). *Monthly Notices of the Royal Astronomical Society: Letters* 471 (1), pp. L131–L134.

Eatough, R., Keane, E., and Lyne, A. (2009). *MNRAS* 395, pp. 410–415.

Edwards, R. T., Hobbs, G. B., and Manchester, R. N. (2006). *Monthly Notices of the Royal Astronomical Society* 372 (4), pp. 1549–1574.

Farr, W. M. et al. (2011). *The Astrophysical Journal* 742 (2), p. 81.

Farrow, N., Zhu, X.-J., and Thrane, E. (2019). *The Astrophysical Journal* 876 (1), p. 18.

Ferdman, R. D. et al. (2010). *The Astrophysical Journal* 711 (2), pp. 764–771.

Ferdman, R. D. et al. (2013). *The Astrophysical Journal* 767, p. 85.

Ferdman, R. D. et al. (2014). *Monthly Notices of the Royal Astronomical Society* 443 (3), pp. 2183–2196.

Ferdman, R. D. et al. (2020). *Nature* 583 (7815), pp. 211–214.

Folkner, W. M. and Park, R. S. (2018). *JPL IOM 392R-18-004*.

Folkner, W. M., Park, R. S., and Jacobson, R. A. (2016). *JPL IOM 392R-16-003*.

Fonseca, E., Stairs, I. H., and Thorsett, S. E. (2014). *The Astrophysical Journal* 787 (1), p. 82.

Fortin, M. et al. (2016). *Astronomy and Astrophysics* 586, A109.

Foster, R. S., Fairhead, L., and Backer, D. C. (1991). *The Astrophysical Journal* 378, p. 687.

Freire, P. C. C. and Ridolfi, A. (2018). *Monthly Notices of the Royal Astronomical Society* 476 (4), pp. 4794–4805.

Freire, P. C. C. and Wex, N. (2010). *Monthly Notices of the Royal Astronomical Society* 409, pp. 199–212.

Freire, P. C. C. et al. (2003). *Monthly Notices of the Royal Astronomical Society* 340 (4), pp. 1359–1374.

Freire, P. C. C. et al. (2012). *Monthly Notices of the Royal Astronomical Society* 423 (4), pp. 3328–3343.

Gehrels, N. et al. (2000). *Nature* 404 (6776), pp. 363–365.

Glebbeek, E. et al. (2013). *Monthly Notices of the Royal Astronomical Society* 434 (4), pp. 3497–3510.

Gold, T. (1968). *Nature* 218, p. 731.

Goodfellow, I., Bengio, Y., and Courville, A. (2016). MIT Press.

Gotthelf, E. V. et al. (2013). *The Astrophysical Journal* 773 (2), p. 141.

Grenier, I. A. (2000). *Astronomy and Astrophysics* 364, pp. L93–L96.

Guinot, B. (1988). *Astronomy and Astrophysics* 192 (1-2), pp. 370–373.

Guyon, I. (1997). *AT & T Bell Laboratories*.

Halpern, J. P. and Gotthelf, E. V. (2010). *The Astrophysical Journal* 709 (1), pp. 436–446.

Hamil, O. et al. (2015). *Physical Review D* 91 (6).

Haniewicz, H. T. (2020). *ascl:2007.007*.

Haniewicz, H. T. et al. (2021). *Monthly Notices of the Royal Astronomical Society* 500, pp. 4620–4627.

Hankins, T. H. (2017). *Proceedings of the International Astronomical Union* 13 (S337), pp. 29–32.

Hankins, T. H. and Rickett, B. J. (1975). *Methods in Computational Physics. Volume 14 - Radio astronomy*. Vol. 14, pp. 55–129.

— (1986). *The Astrophysical Journal* 311, p. 684.

Hastad, J. (1988). *Combinatorica*, pp. 75–81.

Heger, A. et al. (2003). *The Astrophysical Journal* 591 (1), pp. 288–300.

Helfand, D. J., Manchester, R. N., and Taylor, J. H. (1975). *The Astrophysical Journal* 198, pp. 661–670.

Helfand, D. J. and Tademaru, E. (1977). *The Astrophysical Journal* 216, pp. 842–851.

Hessels, J. W. T. et al. (2006). *Science* 311, pp. 1901–1904.

Hewish, A. R. et al. (1968). *Nature* 217.

Hobbs, G. B., Edwards, R. T., and Manchester, R. N. (2006a). *Monthly Notices of the Royal Astronomical Society* 369 (2), pp. 655–672.

Hobbs, G. B., Lyne, A. G., and Kramer, M. (2006b). *Chinese Journal of Astronomy and Astrophysics* 6 (S2), pp. 169–175.

Hobbs, G. et al. (2010). *Classical and Quantum Gravity* 27 (8), p. 084013.

Hoensbroech, A. von, Kijak, J., and Krawczyk, A. (1998).

Hotan, A. W., van Straten, W., and Manchester, R. N. (2004). *Publications of the Astronomical Society of Australia* 21 (3), pp. 302–309.

Hoyle, F., Narlikar, J. V., and Wheeler, J. A. (1964). *Nature* 203, pp. 914–916.

Hulse, R. A. and Taylor, J. H. (1975). *The Astrophysical Journal* 456, pp. L51–L53.

Iglewicz, B. and Hoaglin, D. C. (1993). Milwaukee, WI: ASQC Quality Press, p. 11.

International Astronomical Union (1974). English. Reidel.

Jackson, J. (1998). New York: Wiley.

Jacoby, B. A. et al. (2006). *The Astrophysical Journal Letters* 644 (2), pp. L113–L116.

Janka, H.-T. (2013). *Monthly Notices of the Royal Astronomical Society* 434 (2), pp. 1355–1361.

Janssen, G. H. et al. (2008). *Astronomy and Astrophysics* 490 (2), pp. 753–761.

Jiang, L. et al. (2013). *International Journal of Modern Physics: Conference Series* 23, pp. 95–98.

Kaspi, V. M., Taylor, J. H., and Ryba, M. F. (1994a). *The Astrophysical Journal* 428, p. 713.

Kaspi, V. M. et al. (1994b). *Compact Stars in Binaries*. Ed. by J. van Paradijs, E. P. J. van den Heuvel, and E. Kuulkers. Vol. 165. IAU Symposium, p. 271.

Kaspi, V. M. et al. (2001). *The Astrophysical Journal* 560 (1), pp. 371–377.

Keane, E. F. et al. (2012). *Monthly Notices of the Royal Astronomical Society: Letters* 425 (1), pp. L71–L75.

Keith, M. J. et al. (2009). *Monthly Notices of the Royal Astronomical Society* 393 (2), pp. 623–627.

Keith, M. J. et al. (2010). *Monthly Notices of the Royal Astronomical Society* 409 (2), pp. 619–627.

Kendall, M. (1994). Chichester, West Sussex: John Wiley & Sons.

Kerr, R. et al. (2019). *The Astrophysical Journal* 874 (2), p. 147.

Knuth, D. E. (1976). *SIGACT News* 8 (2), pp. 18–24.

Komesaroff, M. M. (1970). *Nature* 225 (5233), pp. 612–614.

Konacki, M., Wolszczan, A., and Stairs, I. H. (2003). *The Astrophysical Journal* 589 (1), pp. 495–502.

Konar, S. and Chahal, M. (2019). *Research Notes of the AAS* 3 (6), p. 88.

Kramer, M. (1994). *Astronomy and Astrophysics Supplement Series* 107, pp. 527–539.

Kramer, M. et al. (1994). *Astronomy and Astrophysics Supplement Series* 107, pp. 515–526.

Kramer, M. et al. (2006). *Science* 314 (5796), pp. 97–102.

Kramer, M. et al. (2007). *Monthly Notices of the Royal Astronomical Society* 377 (1), pp. 107–119.

Krishnan, V. V. et al. (2020). *Science* 367 (6477), pp. 577–580.

Kwofie, J. A. (2018). MA thesis. University of Manchester.

Lam, M. T. (2017). *ascl:1706.011*.

Lancaster, H. O. (1969). Wiley & Sons.

Lange, C. et al. (2001). *Monthly Notices of the Royal Astronomical Society* 326 (1), pp. 274–282.

Large, M. I., Vaughan, A. E., and Mills, B. Y. (1968). *Nature* 220, p. 340.

Lasky, P. D. et al. (2017). *The Astrophysical Journal* 843 (1), p. L1.

Lazarus, P. et al. (2015). *The Astrophysical Journal* 812 (1), p. 81.

Lazarus, P. et al. (2016). *The Astrophysical Journal* 831 (2), p. 150.

LeCun, Y. et al. (1989). *Neural Computation* 1 (4), pp. 541–551.

Lentati, L. et al. (2017). *Monthly Notices of the Royal Astronomical Society* 468 (2), pp. 1474–1485.

Leys, C. et al. (2013). *Journal of Experimental Social Psychology* 49 (4), pp. 764–766.

Lorimer, D. R. and Kramer, M. (2005). Cambridge University Press.

Lorimer, D. R. et al. (2004). *Monthly Notices of the Royal Astronomical Society* 347 (2), pp. L21–L25.

Lynch, R. S. et al. (2018). *The Astrophysical Journal* 859 (2), p. 93.

Lyne, A. G. and Rickett, B. J. (1968). *Nature* 218, pp. 326–330.

Madison, D. R. et al. (2019). *The Astrophysical Journal* 872 (2), p. 150.

Manchester, R. N. (1971). *Astrophysical Journal Supplement* 23, p. 283.

— (1977). San Francisco: W.H. Freeman.

Manchester, R. N., Durdin, J. M., and Newton, L. M. (1985). *Nature* 313 (6001), pp. 374–376.

Manchester, R. N., Taylor, J. H., and Huguenin, G. R. (1975). *The Astrophysical Journal* 196, pp. 83–102.

Manchester, R. N. et al. (2001). *Monthly Notices of the Royal Astronomical Society* 328 (1), pp. 17–35.

Manchester, R. N. et al. (2005). *The Astronomical Journal* 129 (4), pp. 1993–2006.

Manessi, F. and Rozza, A. (2018). *24th International Conference on Pattern Recognition (ICPR)*, pp. 61–66.

Mardia, K. V. (2000). J. Wiley.

Martinez, J. G. et al. (2015). *The Astrophysical Journal* 812.

Martinez, J. G. et al. (2017). *The Astrophysical Journal Letters* 851 (2), p. L29.

Mazzali, P. A. et al. (2007). *Science* 315 (5813), p. 825.

McLaughlin, M. A. et al. (2005). *Binary Radio Pulsars*. Ed. by F. A. Rasio and I. H. Stairs. Vol. 328. Astronomical Society of the Pacific Conference Series, p. 43.

McLaughlin, M. A. et al. (2006). *Nature* 439 (7078), pp. 817–820.

McMillan, P. J. (2017). *Monthly Notices of the Royal Astronomical Society* 465 (1), pp. 76–94.

McMillan, P. J. and Binney, J. J. (2010). *Monthly Notices of the Royal Astronomical Society* 402 (2), pp. 934–940.

Mignani, R. (2009). *The Messenger* 138, pp. 19–25.

Miyaji, S. et al. (1980). *Publications of the Astronomical Society of Japan* 32, pp. 303–329.

Mueller, H. (1948). *Optical Society of America* 338, p. 661.

Nader, A. and Azar, D. (2020). *Proceedings of the 2020 Genetic and Evolutionary Computation Conference Companion*. GECCO '20. Cancún, Mexico: Association for Computing Machinery, pp. 145–146.

NANOGrav Collaboration et al. (2015). *The Astrophysical Journal* 813 (1), 65, p. 65.

Nocedal, J. and Wright, S. J. (2006). *Numerical Optimization*. New York, NY: Springer New York, pp. 30–65.

Nowakowski, L. A. (1983). *Astronomy and Astrophysics* 127 (2), pp. 259–262.

— (2000). *IAU Colloq. 177: Pulsar Astronomy - 2000 and Beyond*. Ed. by M. Kramer, N. Wex, and R. Wielebinski. Vol. 202. Astronomical Society of the Pacific Conference Series, p. 219.

Ou, Z. W. et al. (2016). *Monthly Notices of the Royal Astronomical Society* 457 (4), pp. 3922–3933.

Øvergård, T. and Østgaard, E. (1991). *Canadian Journal of Physics* 69 (1), pp. 8–15.

Özel, F. and Freire, P. C. C. (2016). *Annual Review of Astronomy and Astrophysics* 54 (1), pp. 401–440.

Özel, F. et al. (2012). *The Astrophysical Journal* 757 (1), p. 55.

Pacini, F. (1968). *Nature* 219, pp. 145–146.

Paczyński, B. (1976). *Structure and Evolution of Close Binary Systems*. Ed. by P. Eggleton, S. Mitton, and J. Whelan. Vol. 73. IAU Symposium, p. 75.

Parent, E. et al. (2018). *The Astrophysical Journal* 861 (1), p. 44.

Parent, E. et al. (2019). *The Astrophysical Journal* 886 (2), p. 148.

Patel, C. et al. (2018). *The Astrophysical Journal* 869 (2), p. 181.

Pedregosa, F. et al. (2011). *Journal of Machine Learning Research* 12, pp. 2825–2830.

Pejcha, O. and Thompson, T. A. (2015). *The Astrophysical Journal* 801 (2), p. 90.

Pejcha, O., Thompson, T. A., and Kochanek, C. S. (2012). *Monthly Notices of the Royal Astronomical Society* 424 (2), pp. 1570–1583.

Pennucci, T. T. (2019). *The Astrophysical Journal* 871 (1), p. 34.

Perera, B. B. P. et al. (2010). *The Astrophysical Journal* 721 (2), pp. 1193–1205.

Perera, B. B. P. et al. (2014). *The Astrophysical Journal* 787 (1), p. 51.

Perrin, F. (1942). *The Journal of Chemical Physics* 10 (7), pp. 415–427.

Petit, G. (2004). *Proceedings of the Journées 2003 "Systèmes deréférence spatio-temporels": Astrometry*, pp. 314–317.

Philippov, A., Timokhin, A., and Spitkovsky, A. (2020). *Physical Review Letters* 124 (24), p. 245101.

Pinheiro, P. O., Collobert, R., and Dollár, P. (2015). *NIPS*.

Pinheiro, P. O. et al. (2016). *ECCV*.

Pletsch, H. J. et al. (2013). *The Astrophysical Journal Letters* 779 (1), p. L11.

Podsiadlowski, P., Rappaport, S., and Pfahl, E. D. (2002). *The Astrophysical Journal* 565 (2), pp. 1107–1133.

Podsiadlowski, P. et al. (2004). *The Astrophysical Journal* 612 (2), pp. 1044–1051.

Postnov, K. A. and Yungelson, L. R. (2014). *Living Reviews in Relativity* 17 (1), p. 3.

Radhakrishnan, V. and Cooke, D. J. (1969). *Astrophysical Letters* 3, pp. 225–229.

Rankin, J. M. (1983). *The Astrophysical Journal* 274, pp. 333–358.

Ransom, S. M. (2001). PhD thesis. Harvard University.

Ridolfi, A. et al. (2019). *Monthly Notices of the Royal Astronomical Society* 490 (3), pp. 3860–3874.

Ritter, H. and King, A. R. (2001). *Evolution of Binary and Multiple Star Systems*. Ed. by P. Podsiadlowski et al. Vol. 229. Astronomical Society of the Pacific Conference Series, p. 423.

Rutherford, E. and Nuttal, J. M. (1913). *Philosophical Magazine* 26 (6), pp. 702–712.

Samuel, A. L. (1959). *IBM J. Res. Dev.* 3, pp. 210–229.

Shalev-Shwartz, S. and Ben-David, S. (2014). New York, NY, USA: Cambridge University Press, pp. 234–235.

Shao, L. et al. (2017). *Physics Review X* 7 (4), p. 041025.

Shapiro, S. L. and Teukolsky, S. A. (1983).

Shibata, M. et al. (2019). *Physics Review D* 100 (2), p. 023015.

Shklovskii, I. S. (1970). *Soviet Astronomy* 13, p. 562.

Sigurdsson, S. (2003). *Science* 301 (5630), pp. 193–196.

Silva, H. O. et al. (2018). *Physics Review Letters* 120 (13), p. 131104.

Smith, S. (1997). San Diego, CA: California Technical Pub.

Smits, R. et al. (2009). *Astronomy and Astrophysics* 493 (3), pp. 1161–1170.

Soglasnov, V. A. (2000). *IAU Colloq. 177: Pulsar Astronomy - 2000 and Beyond*. Ed. by M. Kramer, N. Wex, and R. Wielebinski. Vol. 202. Astronomical Society of the Pacific Conference Series, p. 237.

Solomon, S. C. et al. (2007). *Space Science Reviews* 131, pp. 3–39.

Splaver, E. M. et al. (2002). *The Astrophysical Journal* 581 (1), pp. 509–518.

Staelin, D. H. and Reifenstein, E. C. (1968). *Science* 162, pp. 1481–1483.

Stairs, I. H., Thorsett, S. E., and Camilo, F. (1999). *The Astrophysical Journal Supplement Series* 123 (2), pp. 627–638.

Stehman, S. V. (1997). *Remote Sensing of Environment* 62 (1), pp. 77–89.

Steltner, B. et al. (2021). *The Astrophysical Journal* 909 (1), p. 79.

Stokes, G. G. (1851). *Transactions of the Cambridge Philosophical Society* 9, p. 399.

Stovall, K. (2013). PhD thesis. The University of Texas at San Antonio.

Stovall, K. and et al. (2019). *The Astrophysical Journal* 870, p. 74.

Stovall, K. et al. (2014). *The Astrophysical Journal* 791 (1), p. 67.

Stovall, K. et al. (2018). *The Astrophysical Journal Letters* 854 (2), p. L22.

Sturrock, P. A. (1971). *The Astrophysical Journal* 164, p. 529.

Swiggum, J. K. et al. (2015). *The Astrophysical Journal* 805 (2), 156, p. 156.

Tan, C. M. et al. (2018). *The Astrophysical Journal* 866 (1), p. 54.

Tauris, T. M. and Janka, H.-T. (2019). *The Astrophysical Journal Letters* 886 (1), p. L20.

Tauris, T. M., Langer, N., and Kramer, M. (2011). *Monthly Notices of the Royal Astronomical Society* 416 (3), pp. 2130–2142.

— (2012). *Monthly Notices of the Royal Astronomical Society* 425 (3), pp. 1601–1627.

Tauris, T. M., Langer, N., and Podsiadlowski, P. (2015). *Monthly Notices of the Royal Astronomical Society* 451 (2), pp. 2123–2144.

Tauris, T. M. and van den Heuvel, E. P. J. (2006). Edited by: Lewin and van der Klis. Cambridge University Press.

Tauris, T. M., van den Heuvel, E. P. J., and Savonije, G. J. (2000). *The Astrophysical Journal* 530 (2), pp. L93–L96.

Tauris, T. M. et al. (2013). *The Astrophysical Journal Letters* 778 (2), p. L23.

Tauris, T. M. et al. (2017). *The Astrophysical Journal* 846.

Taylor, J. H. (1987). Ed. by M. A. H. MacCallum. Cambridge University Press, p. 209.

— (1989). *Timing Neutron Stars*. Springer Netherlands, pp. 17–26.

— (1992). *Philosophical Transactions of the Royal Society of London. Series A: Physical and Engineering Sciences* 341 (1660), pp. 117–134.

Taylor, J. H. and Weisberg, J. M. (1989). *The Astrophysical Journal* 345, p. 434.

Taylor, S., Rosten, E., and Drummond, T. (2009). *2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp. 15–22.

Thompson, C., Lyutikov, M., and Kulkarni, S. R. (2002). *The Astrophysical Journal* 574 (1), pp. 332–355.

Thorsett, S. E. and Chakrabarty, D. (1999). *The Astrophysical Journal* 512 (1), pp. 288–299.

Tiburzi, C. et al. (2021). *A&A* 647, A84.

Timokhin, A. N. and Arons, J. (2012). *Monthly Notices of the Royal Astronomical Society* 429 (1), pp. 20–54.

Tsang, D. et al. (2012). *Phys. Rev. Lett.* 108 (1), p. 011102.

Tunaley, J. K. E. (1974). *Journal of Statistical Physics* 11, pp. 397–408.

Turolla, R., Zane, S., and Watts, A. L. (2015). *Reports on Progress in Physics* 78 (11), p. 116901.

van den Heuvel, E. P. J. (2007). *The Multicolored Landscape of Compact Objects and Their Explosive Origins*. Ed. by T. di Salvo et al. Vol. 924. American Institute of Physics Conference Series, pp. 598–606.

van der Walt, S., Colbert, S. C., and Varoquaux, G. (2011). *Computing in Science Engineering* 13 (2), pp. 22–30.

van Leeuwen, J. et al. (2015). *The Astrophysical Journal* 798 (2), p. 118.

van Paradijs, J. and McClintock, J. E. (1995). *X-ray Binaries*. Ed. by J. van Paradijs, W. H. G. Lewin, and E. P. J. van den Heuvel. Cambridge University Press. Chap. 2, pp. 58–125.

van Straten, W. (2013). *The Astrophysical Journal Supplement Series* 204 (1), p. 13.

van Straten, W. et al. (2001). *Nature* 412 (6843), pp. 158–160.

Vapnik, V. N. and Chervonenkis, A. Y. (1971). *Theory of Probability & Its Applications* 16 (2), pp. 264–280.

Villatoro, F. R. (2018). Springer International Publishing, pp. 207–240.

Virtanen, P. et al. (2020). *Nature Methods* 17, pp. 261–272.

Weisberg, J. M. and Huang, Y. (2016). *The Astrophysical Journal* 829 (1), p. 55.

Weisberg, J. M. and Taylor, J. H. (2005). *Binary Radio Pulsars*. Ed. by F. A. Rasio and I. H. Stairs. Vol. 328. Astronomical Society of the Pacific Conference Series, p. 25.

Witze, A. (2020). *Nature*.

Wolszczan, A. and Frail, D. (1992). *Nature* 355, pp. 145–147.

Woosley, S. E. (2019). *The Astrophysical Journal* 878 (1), p. 49.

Yao, J. M., Manchester, R. N., and Wang, N. (2017). *The Astrophysical Journal* 835 (1), p. 29.

Yoon, S.-C., Woosley, S. E., and Langer, N. (2010). *The Astrophysical Journal* 725 (1), pp. 940–954.

You, X. P. et al. (2007). *Monthly Notices of the Royal Astronomical Society* 378 (2), pp. 493–506.

Zha, S. et al. (2019). *The Astrophysical Journal* 886 (1), 22, p. 22.

Zhu, W. W. et al. (2014). *The Astrophysical Journal* 781 (2), p. 117.