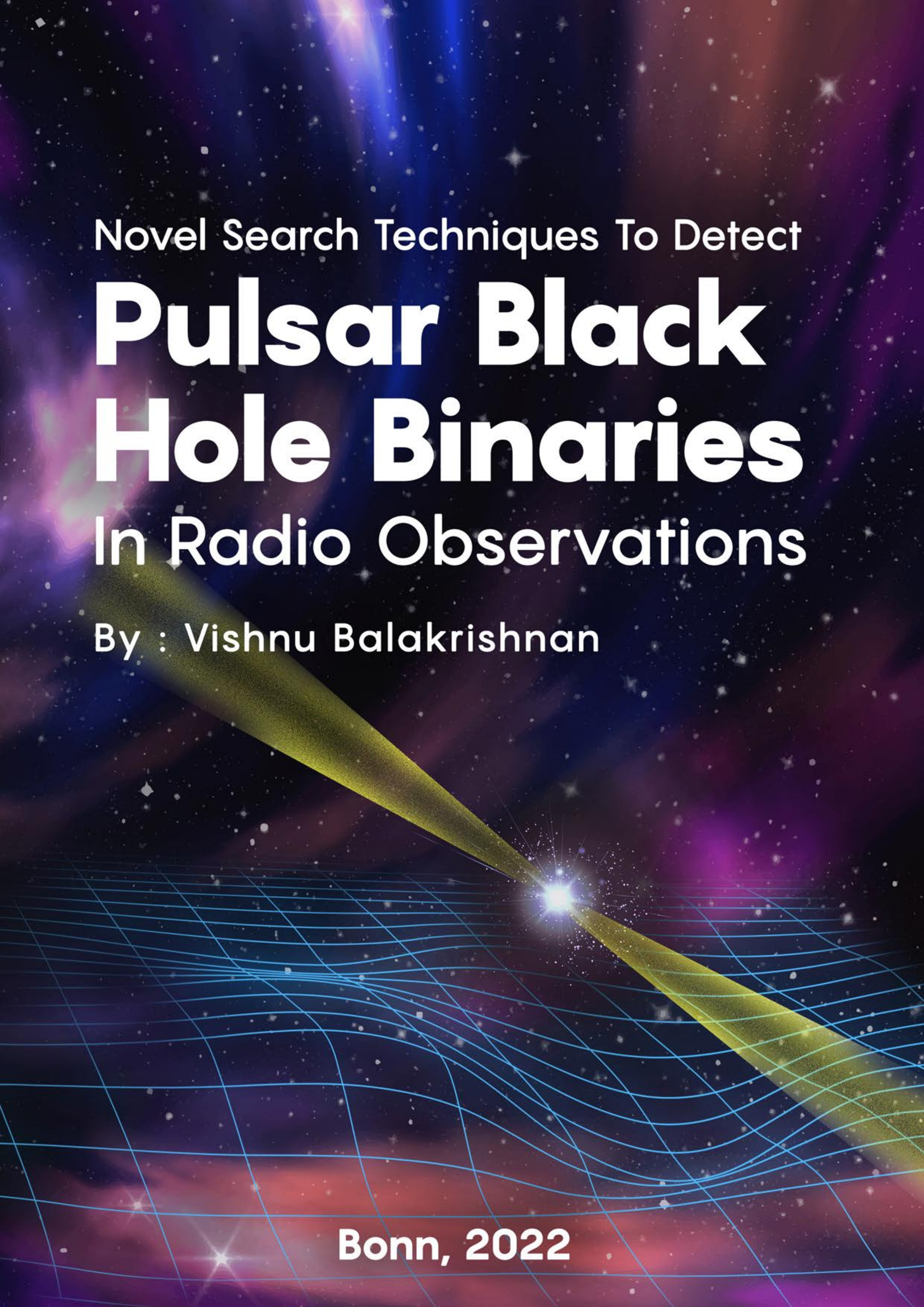


Novel Search Techniques To Detect

Pulsar Black Hole Binaries

In Radio Observations

By : Vishnu Balakrishnan

Bonn, 2022

MAX-PLANCK-INSTITUT FÜR RADIOASTRONOMIE
BONN

Novel Search Techniques to Detect Pulsar Black Hole Binaries in Radio Observations

Dissertation

zur

Erlangung des Doktorgrades (*Dr. rer. nat.*)

der

Rheinischen Friedrich–Wilhelms–Universität, Bonn

vorgelegt von
Vishnu BALAKRISHNAN
aus
Jidhafs, Bahrain

Submitted: 01.10.2021

Bonn

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich–Wilhelms–Universität Bonn

1. Referent: Prof. Dr. Michael Kramer
2. Referent: Prof. Dr. Norbert Langer
Tag der Promotion: 12.01.2022
Erscheinungsjahr: 2022

RHEINISCHE FRIEDRICH–WILHELMS–UNIVERSITÄT BONN

Abstract

by Vishnu Balakrishnan

for the degree of

Doctor rerum naturalium

Pulsars are rapidly rotating highly magnetised neutron stars that emit beams of electromagnetic radiation from their magnetic poles. These compact objects are unique as they are one of the densest forms of matter known in the Universe. Discovering more pulsars are crucial as they have a wide range of scientific applications from studies of strong-field tests of Gravity, neutron star physics, condensed matter physics and cosmology being a few examples. While all pulsars have their own applications, the bulk of this thesis will concentrate on the techniques involved in finding new relativistic binary pulsars particularly the yet to be detected pulsar black-hole binary which can be used to test General Relativity and alternate theories of gravity in the quasi-stationary strong-field regime. In Chapter 1, I lay the foundations of this work by expanding on the basics of pulsar phenomenology, its formation, the diverse types of pulsars already known and their properties in the radio regime. In Chapter 2, I describe the various processes involved in finding a pulsar in a modern pulsar search pipeline. This is followed by a brief history of all the notable past pulsar surveys which provides context to the current generation HTRU-South low latitude (LOWLAT) pulsar survey which is the primary dataset analysed during this PhD thesis. Past searches for Pulsar - Stellar Mass Black Hole (PSR-BH) binaries in LOWLAT assume that the pulsar has a constant acceleration during the course of an observation. However, this assumption breaks down when the observation samples a large fraction of the orbit. This limits the length of search observations for finding compact binaries, and hence their sensitivity.

In Chapter 3, I expand on my comprehensive search for PSR-BH binaries in circular orbits in LOWLAT using the template-bank algorithm and use it to search for recycled and unrecycled PSR-BH binaries in compact orbits. This is currently the most sensitive search for PSR-BH binaries done in Galactic-plane observations in the southern hemisphere. I demonstrate the extra sensitivity factor of 2-2.5 gained from our search compared to previous searches in the same data for PSR-BH binaries with orbital periods in the 6-12 hours range. I also give details about a new GPU pipeline that was developed during this PhD which accelerated our search analysis. Additionally, I give details about our 20 new pulsar discoveries including a new millisecond pulsar J1743−24 which is a rare intermediate spin-period pulsar in a 70.7 day orbit around a light companion star.

I also present updated timing solutions of PSR J1753−2819 - a pulsar similar to PSR J1743−24 but in a much shorter orbit of 9.3 hours and it's likely that the formation of both these binary pulsars cannot be explained by standard binary stellar evolutionary models. I conclude this chapter by using our non-detections of PSR-BH binaries to place limits on short orbital period PSR-BH binaries near the Galactic-Plane ($|b| < 3.5^\circ$). Our results indicate that the existence of nearby ($d \leq 1$kpc) PSR-BH binaries with circular orbits and orbital period range of 4-24 hours is highly unlikely. The possibility of PSR-BH binaries having significantly eccentric orbits or circular orbits shorter than 4 hours cannot currently be ruled out due to them being outside our search range.

In Chapter 4, I describe a novel Machine-learning (ML) pulsar candidate classifier using Semi-Supervised Generative Adversarial Networks (SGAN) which achieved better classification performance than the standard supervised algorithms commonly used in literature using majority unlabelled datasets. This is the first implementation of a Semi-Supervised ML classifier for pulsar candidate classification in literature. The SGAN pipeline achieved an accuracy and mean F-Score of 94.9% trained on only 100 labelled candidates and 5000 unlabelled candidates compared to our standard supervised baseline which scored at 81.1% and 82.7% respectively. This pipeline played a pivotal role in reducing the number of pulsar candidates that needed to be inspected by eye which aided us in finding our earlier mentioned pulsar discoveries. We also describe in detail why this would be a promising solution in the early stages of future pulsar surveys when limited labelled data is available.

In Chapter 5, I push the envelope of sensitivity that can be regained from a binary search pipeline further by describing the first fully coherent radio pulsar search pipeline that can search across all Five Keplerian Parameters. I compare the performance of this pipeline to standard techniques used in literature like acceleration and jerk searches and describe the feasibility of this approach for targeted and blind pulsar surveys in the near future. I demonstrate that a five-parameter search for pulsars with spin-period $P_{\rm spin} \geq 10\,$ms orbiting intermediate mass black holes (IMBH: $M \sim 10^2$ - 10^4M$_\odot$) in Globular clusters with orbital periods in the $P_{\rm orb} = 5 - 10\,T_{\rm obs}$ regime is feasible for observations shorter than 2 hours with an eccentricity limit of 0.1 in the template bank. This is a region in the binary orbital phase-space that cannot be fully explored by other search techniques. Finally in Chapter 6, I summarise the relevant findings of this thesis, and describe some possible future research paths that can be undertaken based on my results.

To my parents, Vijaya and Balakrishnan

Each piece, or part, of the whole nature is always an approximation to the complete truth, or the complete truth so far as we know it. In fact, everything we know is only some kind of approximation, because we know that we do not know all the laws as yet. Therefore, things must be learned only to be unlearned again or, more likely, to be corrected. The test of all knowledge is experiment. Experiment is the sole judge of scientific “truth”.

Richard Feynman
The Feynman Lectures, Introduction

Acknowledgements

Like most members of the scientific community, working on problems related to astronomy for a living has been the fulfillment of a childhood dream of mine. Looking back now and evaluating the entire journey, all I can say is that it has been a long road, filled with incredibly rewarding experiences and seemingly insurmountable challenges along the way. This PhD thesis marks the end of my latest leg of this journey. The title of Doctor rerum naturalium (Latin for ‘doctor of natural sciences’) in Germany is awarded to the individual student who performed the main research presented in a thesis. However, new science is rarely the fruits of individual labour and there are usually several people who have either directly or indirectly aided the student in carrying out their research. I hope I do justice to all the incredible people who have helped me along the way to complete this PhD, and if you are not mentioned here, and if you know me well you would know that it was most likely just a result of my poor recollection skills rather than my design. I would like to first thank my advisors Prof. Michael Kramer and Dr. David Champion. One of the nicest aspects about working with them is the incredible level of trust and freedom they provide in exploring your own ideas. One of the unfortunate drawbacks of doing risky exploratory science is its high probability of failure and during such scenarios, having advisors like them who support you during your successes and failures is invaluable. As the famous poet Rudyard Kipling once wrote, “If you can meet with triumph and disaster And treat those two imposters just the same". I am glad that I had advisors who treated me the same during the ups and downs of my PhD project. Thanks to both of them for going through several earlier drafts of this PhD thesis and giving me constructive feedback and for several helpful suggestions during my Thesis Advisory Committee meetings. Additionally to Prof. Michael Kramer, thank you for sending those weekly group emails during the Covid-19 pandemic. Reading those during these troubled times did make me feel better and gave me a feeling of belonging to this amazing pulsar group you have built here in Bonn. Special thanks for all the efforts you and other members of the HTRU collaboration had to put in the planning, commissioning, observing and storing of Lowlat observations which was required to do the work presented in this thesis. Additionally, thanks for finding the funding to install a new GPU computing cluster hercules2 which has enchanced our computing power and will no doubt aid in finding several new and exciting pulsar discoveries in the near future. I would also like to thank Dr. Ewan Barr who is basically my technical Guru and in my opinion easily one of the best pulsar searchers in the world. I hope my softwares will eventually get the same usability as the ones you have written over the past few years. I would also like to thank Gregory, Ralph, Vivek, Alessandro and Aditya, my next set of technical experts that I turn towards to when I am stuck at a difficult problem. Special thanks to Matthew for making my stay in Melbourne memorable and the several helpful discussions we had including giving me access to the amazing computing cluster ozstar. My three-week visit in Swinburne and drinking the amazing coffee you guys have over there were definitely the highlight

of my PhD. Additionally, thanks to Dr. George Hobbs and Dr. Andrew Cameron for hosting me at the CSIRO institute in Sydney and giving me an opportunity to present a colloquium there. Thanks to Paulo for teaching me how to solve and time a pulsar and Ramesh for helping out with my initial set of Effelsberg observations. I would also like to thank Hans, Henning, Jens, Maude and Prajwal for inviting me to their badminton group and going easy on me during rallies. To my office-mate Prajwal, Rafa will remain the Goat in our hearts, hopefully he becomes the Goat in the total Grandslam count and the head to head record against other members of the Big 3. People often go to Paris to find "true" love, I am glad that we did find Rafa-our one true love there! Also cheers to the many memories we had at the SAP headquarters in Walldorf, the SAP meeting we had at Effelsberg and our trip to Spain. To Madhuri, you have been a special friend of mine during my PhD. Thank you for our coffee sessions, always being my cheerleader and supporting friend. To Shalini, my other great friend in Bonn. Thanks for the several passionate debates we have had over the years which have unfortunately yielded a situation now in which it takes quite some time before we find our first point of disagreement and divergence of view points. You have influenced my thinking in several social challenges that women face in academia and beyond. I have no doubt that you will go on to become a fine astronomer but you are an even better friend. Special regards to some of my other friends in Bonn and in no particular order Joseph, Devina, Lorenz, Ayushi, Manali, Vyoma, Jan, Jaswanth, Surya, Rittika and awesome colleagues Mary, Eleni, Maja, Natasha, Sonia, Joey, Tasha, Richa, Jackie and Weiwei. Cheers to friends back home Jeffin, Shephine, Sruthy and my confidant and friend for life - Vidya. Special thanks to Jason and Ann for the several tennis discussions before, during and after a match between Novak, Roger and Rafa. I would like to also thank all the administrative staff from MPIfR like Kira and Le Tran among others who helped me with all the bureaucratic hurdles involved in living in Germany. I am most grateful to my Mom for always being there for me and supporting me in all my endeavours. I would also like to thank Ajay Chettan, Devi Chechi, Abhi and Diya who are basically equivalent to family. Thank you for hosting me at your house during my undergraduate days. Finally, I would like to thank my great predecessors in pulsar astronomy. All the way from our pioneers Dr. Baade, Dr. Zwicky, Dr. Gold, Dr. Hewish and Dr. Burnell to all the amazing professors, postdocs and students around the world today. I have learnt a lot from your works. What the last four years has taught me is that pulsar astronomy is a broad and rich field, and I have only now begun to scratch the surface. The only reason we are making new progress in this field, is because we are 'Standing on the shoulders of giants'. Thank you for that, and below in this thesis you can find my leaf to the metaphorical tree of knowledge in pulsar astronomy. I hope this would be useful for someone someday.

Contents

List of Figures

List of Tables

Nomenclature

Frequently Used Symbols

b	Galactic latitude
B	Magnetic flux density
B_0	Magnetic flux density at the surface
g	Gravitational acceleration
G	Gravitational constant
c	Speed of light
δ	Declination
e	Eccentricity (or electron charge or numerical constant)
$E_{\mathrm{B-V}}$	Colour excess
Z	Metallicity (or atomic number)
h	Planck's constant
H	Hour angle
i	Inclination
J	Orbital angular momentum
k	Coulomb's constant
k_{B}	Boltzmann's constant
σ	Stefan–Boltzmann constant
λ	Longitude
L	Luminosity
μ	Mean molecular weight
m	Mass (or apparent magnitude)
M	Absolute magnitude (or mass)
ν	Frequency
n	Numerical density (or neutron)
π	Parallax (or numerical constant)
p	Momentum (or proton)
P	Pressure (or period)
q	Mass ratio
r	Distance
R	Radius
t	Time
T	Temperature
ϕ	Latitude (or angle)
v	Speed
V	Volume
X	Fraction of hydrogen

Y	Fraction of helium
Ω	Solid angle
ω	Periastron (or angular frequency)

Numerical Constants

π $= 3.14156$

1 rad $= 57.296$ degrees

e $= 2.7183$

$\log e$ $= 0.4343 = \ln(10)^{-1}$

Physical Constants

Speed of light	$c = 2.9979 \times 10^{10}\,\mathrm{cm\,sec^{-1}}$
Gravitational constant	$G = 6.670 \times 10^{-8}\,\mathrm{dynes\,cm^2\,gr{-}1}$
Planck's constant	$h = 6.626 \times 10^{-27}\,\mathrm{erg\,sec}$
Coulomb's constant	$k = 1$
Boltzmann's constant	$k_{\mathrm{B}} = 1.381 \times 10^{-16}\,\mathrm{erg\,deg^{-1}}$
Stefan–Boltzmann constant	$\sigma = 5.6704 \times 10^{-5}\,\mathrm{erg\,cm^{-2}\,s^{-1}\,K^{-4}}$
Electron mass	$m_{\mathrm{e}} = 9.110 \times 10^{-28}\,\mathrm{gr}$
Proton mass	$m_{\mathrm{p}} = 1.673 \times 10^{-24}\,\mathrm{gr}$

Astronomical Constants

Astronomical unit (1 AU) $= 1.496 \times 10^{13}$ cm

Parcec (1 pc) $= 3.086 \times 10^{18}$ cm

Julian light year (1 ly) $= 9.460730472 \times 10^{17}$ cm

Julian year (1 yr) $= 3.15576 \times 10^{7}$ sec

Solar mass ($1\,\mathrm{M}_\odot$) $= 1.989 \times 10^{27}$ gr

Solar radius ($1\,\mathrm{R}_\odot$) $= 6.960 \times 10^{10}$ cm

Solar luminosity ($1\,\mathrm{L}_\odot$) $= 3.9 \times 10^{33}\,\mathrm{erg\,s^{-1}}$

Absolute magnitude of the Sun (M_{V}) $= 4.77$

Apparent magnitude of the Sun (m_{V}) $= -26.7$

Effective temperature of the Sun (T_{eff}) $= 5770$ K

Chapter 1

An Introduction to Radio Pulsars

> With all reserve we advance the view that a super-nova represents the transition of an ordinary star into a neutron star, consisting mainly of neutrons. Such a star may possess a very small radius and an extremely high density.
>
> Baade and Zwicky, PNAS May 1, 1934 20 (5) 259-263

Contents

1.1 Historical Context

1.1.1 Neutron Star Hypothesis

The neutron star hypothesis was first put forward by Baade & Zwicky (1934) in order to explain the tremendous amount of energy required to power a supernova and to simultaneously produce energetic cosmic rays. This remained a theoretical prediction for a long time until 33 years later Jocelyn Bell discovered a repeating radio signal in observations taken at the Mullard Radio Astronomy Observatory at Cambridge (Hewish et al., 1968). The radio signal was observed to follow sidereal time[1] and repeated every 1.337 seconds. The absence of parallax indicated that the source must lie outside the solar system and the observed frequency sweep of the signal was used to infer that these signals originated from local objects at a galactic distance scale. The extremely rapid spin-period of the radio source along with the instantaneous bandwidth and frequency sweep of the signal was used to calculate an upper-limit for the physical dimension of the source - 4800 km (Hewish et al., 1968). Prior to this discovery, Pacini (1967) hypothesised that neutron stars must be formed after the supernovae explosion (at least in some cases) and further estimated that the electromagnetic radiation emitted by it should give an outward momentum to accelerate the expansion of supernova remnants after its initial explosion. There was good reason to believe this as if you take the size of the Crab Nebula and divide it by its measured expansion rate, this number was an under-estimation of its total lifetime. Post discovery, Hewish et al. (1968) favoured an oscillating white dwarf (WD) model whereas Gold (1968) hypothesised that rotating neutron stars could be the origin of this pulsating source. Around this time, a daily telegraph reporter Anthony R. Michaelis coined the term ‘pulsar’ (a portmanteau of ‘pulsating star’) to describe this discovery. The rotating neutron star hypothesis further gained credence with the discovery of pulsars in supernova remnants like Vela (Large et al., 1968) which had a spin-period of $\sim$ 89 ms and in particular the Crab Pulsar (Staelin & Reifenstein, 1968) initially named NP 0527. Lovelace et al. (1968) measured the spin-period of NP 0532 to be $\sim$ 33 ms thereby making the oscillating WD model extremely unlikely as such accurate and rapid spin-periods require a compact object with extremely high density beyond what is theoretically possible with an oscillating WD model.

1.2 Formation

The core of a star for most of its lifetime is in a finely tuned ‘hydrostatic equilibrium’. The outward thermal pressure caused due to the energy generated from nuclear fusion reactions is balanced by the inward gravitational force due to the mass of the star. For high-mass progenitor stars ($m > 8\mathrm{M}_\odot$), over time, the star fuses heavier elements starting from hydrogen and helium until it reaches a core of iron and nickel. The fusion of iron and nickel is an *endothermic* reaction, therefore no further reaction takes place

---

[1] This is a timescale system used in astronomy based on Earth rotation rate measured relative to fixed stars.

forcing the star to undergo a gravitational collapse. If the mass of the core exceeds the Chandrasekhar limit of about $1.4\mathrm{M}_{\odot}$ (Chandrasekhar, 1931) but is below the Tolman-Oppenheimer-Volkoff (TOV) (Oppenheimer & Volkoff, 1939) which is an upper bound for the mass of neutron star i.e. gravity is balanced by neutron degeneracy pressure, then a neutron star is left behind after a core-collapse supernova explosion which blows away the outer region of the star. The neutron star retains most of the angular momentum of its progenitor star and due its now reduced radius, the moment of inertia of the star drops rapidly. Due, to the conservation of angular momentum, this leads to a increase in the spin-frequency of the pulsar. If the star collapses in a spherically symmetric way, Woltjer (1964) first suggested that due to conservation of magnetic flux ($B \propto R^{-2} \propto \rho^{\frac{2}{3}}$), the initial magnetic field of the progenitor star gets amplified to values of up to 10^{14-16}G. Neutron stars are the densest objects known in the universe with an average mass density $\langle\rho\rangle \sim 6.7 \times 10^{17}\,\mathrm{kgm}^{-3}$. The behaviour of matter in these extreme densities is not well known as such environments cannot be replicated in terrestrial laboratories. Therefore, pulsars offer a unique tool to explore cold nuclear matter at densities higher than that of the atomic nucleus, a major unsolved problem in modern physics (Özel & Freire, 2016). The physical properties of a neutron star can be understood if we know the *equation of state* i.e the relation between density and pressure. This could for example help us find the relation between mass and radius of a neutron star. The maximum mass for a neutron star is of particular interest as it can place constraints on the equation of state. PSR J0348+0423 with a mass of $2.01{\pm}0.04\mathrm{M}_{\odot}$ (Antoniadis et al., 2013) and PSR J0740+6620 with a mass range at the 68.3% credibility interval of $2.08^{+0.07}_{-0.07}$ $\mathrm{M}_{\odot}$ (Fonseca et al., 2021) display the highest known neutron star mass measurements till date. Most models predict a maximum mass of $2\mathrm{M}_{\odot}$ (Lattimer & Prakash, 2001), which can increase depending on the magnetic field (Cardall et al., 2001) and rotating rate. The mass of a neutron star can be measured accurately through pulsar timing (see section 1.7). A recent review highlighting the range of measured masses of neutron stars can be found in Özel & Freire (2016) which shows a range between $\sim 1.1 - 2.08\mathrm{M}_{\odot}$ where we have updated the upper limit based on the measured mass of PSR J0740+6620. The characteristic mass of a pulsar which is used often for our calculations in this work is $1.4\mathrm{M}_{\odot}$. Measuring the radius of the neutron star is more difficult. One technique is to measure the radius of the neutron star through spectroscopy by measuring the thermal emission at X-ray frequencies which is used to measure the apparent angular size of the star. An alternative is to measure the effect of the neutron-star spacetime on the emission and extract the radius. Current measurements indicate values around 10-12 km (Lattimer & Prakash, 2001). Recently, the gravitational wave event of a binary neutron star merger GW170817 was used to measure the radius of the two neutron stars (Abbot et al., 2018). Additionally, X-ray observations using Neutron Star Interior Composition Explorer (NICER) were also used to constrain the radius of the pulsar PSR J0030+0451 to $13.02^{+1.24}_{-1.06}$ km (Miller et al., 2019) and PSR J0740+6620 to $12.39^{+1.30}_{-0.98}$ km (Riley et al., 2021). Radius measurement of PSR J0740+6620 was obtained based on informative priors from radio timing by the results of Fonseca et al. (2021). We refer the readers to Özel & Freire (2016) and references within for an in-depth review of the different techniques currently in use to measure the radius of a neutron star.

1.3 The lighthouse model

The canonical model used to explain the pulsar mechanism is the light house model. An example is shown in figure 1.1. In this model, the neutron star is assumed to rotate in a purely dipolar magnetic field which is inclined with respect to the rotation axis of the pulsar with a misalignment angle α. This misalignment is what causes the signal to pulsate as viewed from Earth when it crosses our line of sight similar to a light-house. This rotating magnetic field B induces an electric field E outside the pulsar surface. Electrons at the surface of the neutron star gain energy from this external E-field and they subsequently get converted into plasma. The plasma-filled surrounding dominated by the B-field is called the magnetosphere of the pulsar. The magnetosphere can be thought of as an extension of the neutron star as it experiences the same E x B force and co-rotates with the pulsar. This co-rotation can only be maintained up-to a certain radius where the velocity of the plasma approaches the speed of light. This is usually called the radius of the light cylinder[2] and is calculated as:

$$R_{LC} = \frac{c}{\Omega} = \frac{cP}{2\pi} \sim 4.77 \times 10^4 \mathrm{km} \left(\frac{\mathrm{P}}{\mathrm{s}}\right), \tag{1.1}$$

where P is the Spin-Period of the Pulsar. The light cylinder divides the magnetic field into two regions ‘closed field lines’ (confined within the light cylinder) and ‘open field lines’ (particles can flow outside the magnetosphere). When charged particles are accelerated in a magnetic field they move along the magnetic field lines. Therefore, the open field lines are likely to be related to the beamed coherent broadband radio emission which we observe. The mechanism underlying the radio emission is not well understood. We refer the readers to Sturrock (1971); Chen & Ruderman (1993); Graham-Smith (2003) and references within for a more in-depth review of this topic.

1.4 Spin Evolution

The spin-period of a pulsar increases over time ($\dot{P} = dP/dt > 0$). This is called *spin-down* and it happens as the pulsar loses rotational kinetic energy through a number of different physical processes. The bulk of the energy is converted into magnetic dipole radiation and a pulsar wind[3]. A tiny portion of this energy loss is also responsible for the radio emission that we typically observe. This energy loss can be calculated (see

---

[2] Defined as a hypothetical region where an object co-rotating with the pulsar is moving at the speed of light

[3] Pulsars as they rotate are believed to power a magnetized particle wind where charged particles are accelerated to relativistic speeds forming a nebula around the pulsar called pulsar wind nebula. An example of this is the Crab Nebula which radiates synchrotron emission across the electromagnetic spectrum.

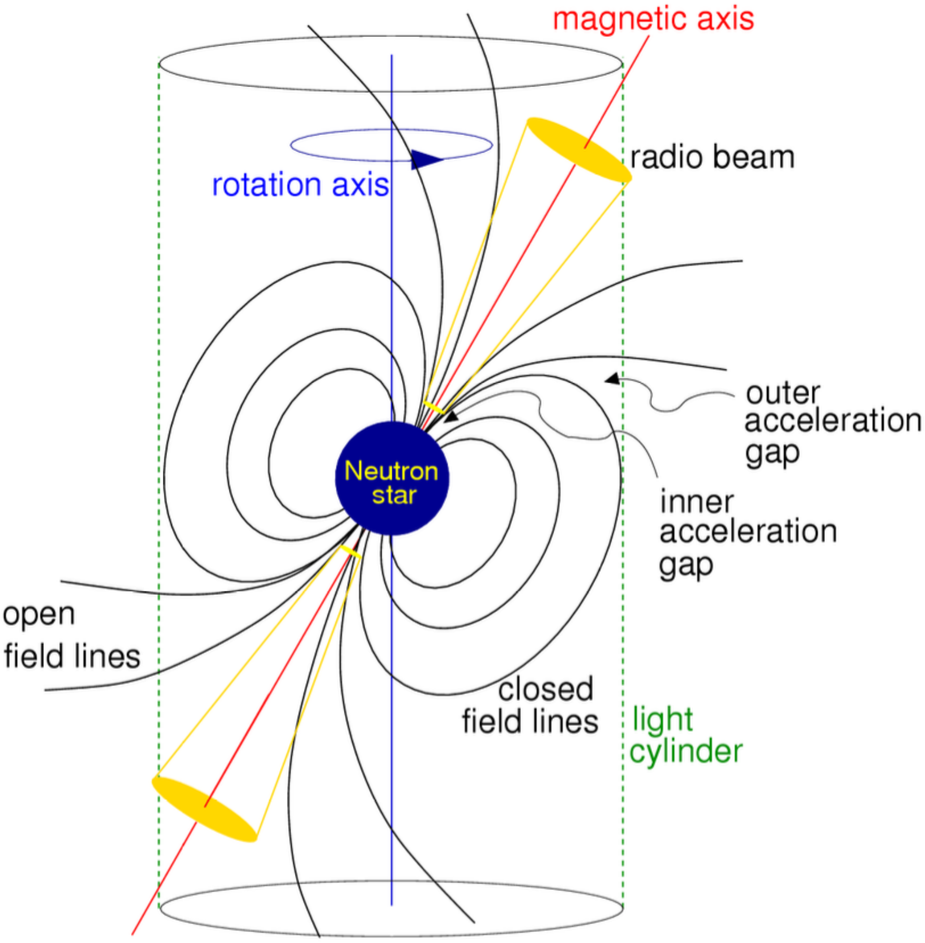


Figure 1.1: An illustration of the toy light-house model used to explain a rotating neutron star and its magnetosphere (not drawn to scale). Figure taken from Lorimer & Kramer (2012).

for example Lorimer & Kramer 2012) as:

$$\dot{E} = \frac{-dE_{rot}}{dt} = \frac{dI\Omega^2/2}{dt} = -I\Omega\dot{\Omega} = 4\pi^2 I\dot{P}P^{-3}, \tag{1.2}$$

where I is the moment of inertia of the pulsar, which is calculated as:

$$I = kM_{\rm pulsar}R^2. \tag{1.3}$$

The value of k depends on the equation of state. For most practical calculations, the neutron star is assumed to be a sphere with uniform density which gives $k = 0.4$, $M_{\rm pulsar} = 1.4M_\odot$ and $R = 10\,\rm km$ which gives a value for $I = 10^{45}\,\rm g\,cm^2$. Substituting these numbers back into equation 1.2 gives:

$$\dot{E} \sim 3.95\times10^{31}{\rm ergs}^{-1}\left(\frac{\dot{\rm P}}{10^{-15}}\right)\left(\frac{\rm P}{\rm s}\right)^{-3}. \tag{1.4}$$

From classical electrodynamics (see for example Jackson 1962), we known that a rotating magnetic dipole with magnetic moment $|m|$ radiates electromagnetic waves at its rotation frequency. The radiation power loss can be written as

$$\dot{E_{\rm dipole}} = \frac{2}{3c^3}|m|^2\Omega^4\sin^2(\alpha). \tag{1.5}$$

If we assume spin-down luminosity is dominated by magnetic dipole radiation loss, then equation 1.5 can be substituted in equation 1.2 to calculate the spin-down of the spin period as:

$$\dot{\Omega} = -\left(\frac{2|m|^2\sin^2\alpha}{3Ic^3}\right)\Omega^3. \tag{1.6}$$

Equation 1.6 can be written in terms of the rotational frequency of the pulsar $\nu = 1/P$ and be expressed as a power law:

$$\dot{\nu} = -K\nu^n, \tag{1.7}$$

where K is a proportionality constant and n is the *braking index*. If we differentiate and re-arrange equation 1.7, the braking index can be calculated as:

$$n = \frac{\nu\ddot{\nu}}{\dot{\nu}^2}. \tag{1.8}$$

Therefore calculating the braking index requires measuring second derivative of the spin-frequency which is only possible for young pulsars that have high spin-down rates. However, in practice the impact of timing irregularities such as glitches (Hobbs et al., 2010; Espinoza et al., 2011) can limit the precision of our measurements. For a pure magnetic dipole radiation, we expect $n = 3$. However, in practise other mechanisms like pulsar wind changes the value of n. For a few pulsars, the braking index have been measured to be ranging from $n = 1.4 - 2.9$ (see for example Kaspi & Helfand 2002). These results show that the assumption of pure dipole radiation is not correct. However, this is a useful approximation in order to characteristic various properties of radio pulsars. For example, by integrating equation 1.7 after writing it in terms of spin period P, and assuming that the spin period of the pulsar at birth is much shorter than the current value ($P_0 \ll P$), we can compute the *characteristic age* of the pulsar. This is given as:

$$\tau_c = \frac{P}{2\dot{P}} \sim 15.8\mathrm{Myr}\left(\frac{\mathrm{P}}{\mathrm{s}}\right)\left(\frac{\dot{\mathrm{P}}}{10^{-15}}\right)^{-1}. \tag{1.9}$$

As mentioned earlier, this is only an order of magnitude estimate due to our incorrect assumption of pure magnetic dipole braking and negligible initial spin-period. This is particularly applicable for MSPs **(see section 1.6.3)** which follow a different evolutionary track and is spun-up to short spin-frequencies by recycling which breaks the assumption of negligible initial spin-period (e.g. Tauris et al. 2012). Another useful quantity, that can be calculated using the above assumptions is the surface magnetic field of the neutron star $\mathrm{B_{surf}}$. For a canonical neutron star with $I = 10^{45}\,\mathrm{g\,cm^2}$, $\alpha = 90°$, $R = 10\,\mathrm{km}$, Lorimer & Kramer (2012) showed that this takes the form:

$$B_{\mathrm{surf}} = 3.2 \times 10^{19} G\sqrt{P\dot{P}} \sim 10^{12} G\left(\frac{\dot{P}}{10^{-15}}\right)^{1/2}\left(\frac{P}{s}\right)^{1/2}. \tag{1.10}$$

If we apply equation 1.10, to the general population of normal pulsars **(see section 1.6.2)**, we get values in the range of 10^{11-13} G. These values agree well with more reliable measurements of B_{surf} by detecting cyclotron radiation in the X-ray spectra of isolated and binary neutron stars (Bignami et al., 2003; Ebbens, 2020).

1.5 Propagation Effects

The radio emission from pulsars, before arriving at our receivers at the telescope, passes through the Interstellar Medium (ISM). This creates various propagation effects that needs to be taken into account during our data analysis. In this section, I will describe some of the most commonly encountered propagation effects for radio pulsar observations.

1.5.1 Dispersion

If space was a perfect vacuum then the broadband emission from pulsars would arrive at Earth simultaneously. However, the intervening ISM consists of cold ionised plasma which causes a delay in the propagation of light with the delay being a function of the radio frequency and mass of the charged particle. This causes the lower frequency components of the signals to be delayed compared to their higher frequency counterparts. The delay is inversely proportional to the mass of the particle, therefore most of the effects of dispersion is dominated by the presence of electrons. This effect was already noticed in the first set of pulsar discoveries by Hewish et al. (1968). An example of the dispersive delay caused in a pulsar observation is shown in figure 1.2. Additionally, the amount of delay is also proportional to a quantity called the dispersion measure (DM) which is proportional to the integrated column density of electrons between the pulsar and the observer. This can be calculated as:

$$\mathrm{DM} = \int_0^d n_e \, dl \, \mathrm{cm}^{-3}\mathrm{pc}, \tag{1.11}$$

where d is the distance to the pulsar, n_e is electron number density which for the ISM is typically around $n_e \simeq 0.03\mathrm{cm}^{-3}$ (Ables & Manchester, 1976). The delay between two different frequency components of a signal can be computed as:

$$\Delta t \simeq 4.15 \times 10^6 \, \mathrm{ms} \times (f_1^{-2} - f_2^{-2}) \times \mathrm{DM} \tag{1.12}$$

We refer the readers to Lorimer & Kramer (2012) for the full derivation of this expression. From equation 1.12, we see that if we measure the pulse arrival time at two different frequencies f_1 and f_2, then we can calculate the DM and substitute it back into equation 1.11. If we then assume an electron density model to get a value for n_e, then by numerically integrating this expression, we can calculate the distance to the pulsar. The two most commonly used electron density models in literature today are the 'NE2001' model (Cordes & Lazio, 2002) and the YMW16 model (Yao et al., 2017). These models were built based on independent measurements of distances via parallax measurements through pulsar timing and/or interferometer techniques and HI absorption line measurements combined with a kinematic rotation model of the Galaxy. Additionally, these models also take into account inhomogeneities of the ISM like the disk and spiral arms of the Galaxy, supernova remnants where the electron density is expected to be higher. In section 2.2.2, I will describe a technique called *dedispersion* to correct for this effect.

1.5.2 Scattering

A second propagation effect is interstellar scattering. This happens when the initial spatially coherent pulsar radiation gets scattered into different line of sights because of the inhomogeneities and turbulence of the ISM. These inhomogeneities can be thought of as scattering disks of different radii which causes a phase shift in the signal due to their different refractive indices. When observed from Earth, this effectively delays parts of the pulse profile of the signal creating the characteristic exponential tailed pulse profile at low frequencies. An example of this phenomenon is shown in figure 1.3. This can be modelled as a convolution of the true pulse profile with a exponential function.

$$I(t) \propto exp(-t/t_s), \tag{1.13}$$

where t_s is the scattering time scale i.e signals emitted at the same time from the pulsar arrives at the detector over a time t_s. The amount of scattering is directly proportional

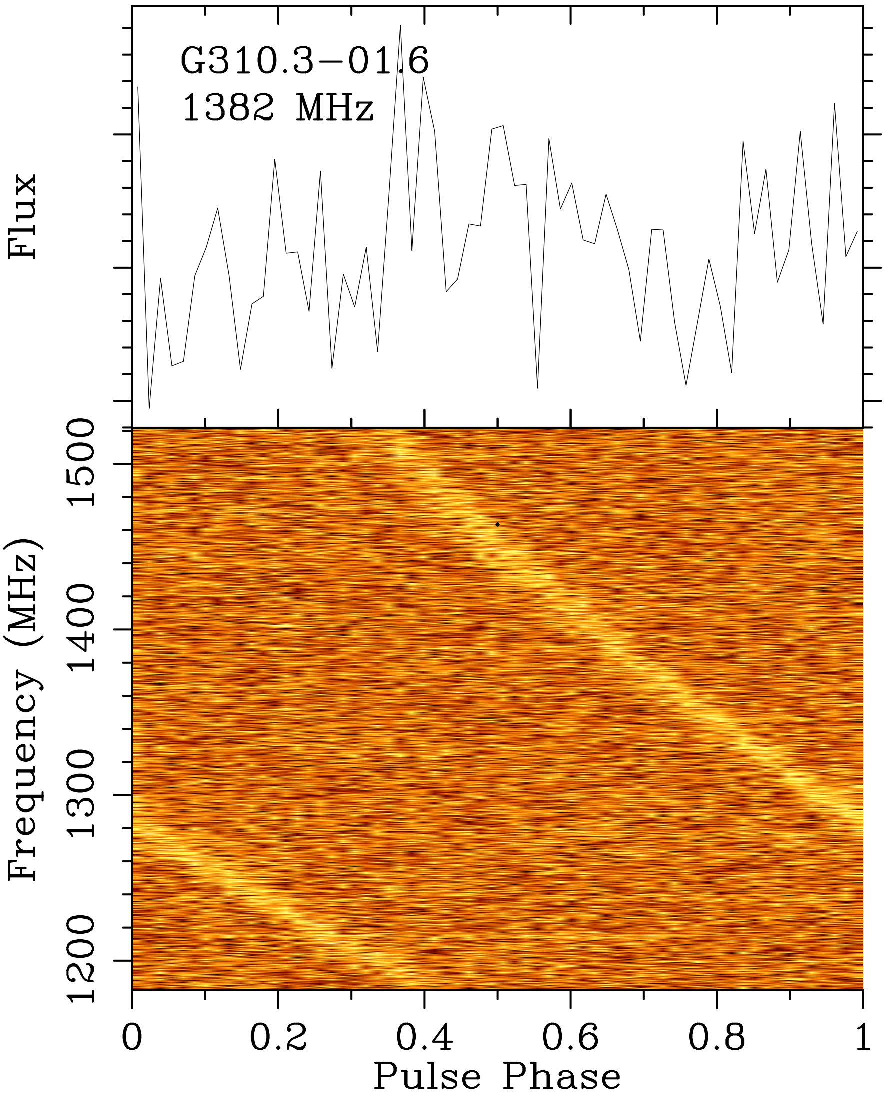


Figure 1.2: An example showing the effect of dispersive delay for PSR J1357-62 in an HTRU-South Low Latitude observation. The bottom plot shows frequency versus rotational phase of a pulsar signal. We can see in this that signals at the lower end of the frequency spectrum arrive much later compared to higher frequency signals and they follow a quadratic relation. On the top, we show the pulse profile which has been completely smeared out. This hinders detection by reducing the final S/N. In section 2.2.2, I will describe a technique called incoherent dedispersion which is a standard technique used for accounting for this in a typical pulsar search pipeline.

to the size of the scattering disk. The most commonly used model to explain this is the thin screen model (Scheuer, 1968), where we assume that all the inhomogeneities in the ISM follow a Kolmogorov spectrum with a thin plasma screen of irregularities situated between the pulsar and the observer. Using this model, $t_s \propto f^{-4}$. Therefore, to minimise the effect of scattering, observing at higher frequencies is more desirable. Additionally, e.g. Bhat et al. (2004) showed by measuring pulse broadening time scales for a wide variety of pulsars that pulsars at high DM are more likely to be scattered.

1.5.3 Scintillation

Another propagation effect similar to scattering is interstellar scintillation. Scintillation is the sudden change in intensity of the pulsar signal caused by the relative motion of the pulsar, the scattering medium and the observer. The irregularities and turbulence in the scattering medium adds phase modulations to the propagating radio waves. These intensity changes have been observed at various time scales and bandwidths. They are analogous to the effect of *twinkling* of stars in the night-sky. This effect was first observed by Lyne & Rickett (1968). Similar to scattering, scintillation is typically modelled assuming a thin screen model (Scheuer, 1968) and has a characteristic frequency dependence known as *scintillation bandwidth* $\Delta f \propto f^4$, where f is the observing frequency. These sudden changes in intensity known as *scintils* are known to be more prominent for nearby pulsars at a low DM Sutton (1971); Backer (1975).

1.6 The Pulsar Family: A tale of diversity

There are currently at least 3176 pulsars known[4]. Naturally, not all pulsars are the same. Around 232 pulsars[5] are known to be located in Globular Clusters (GCs) and a total of 31 (8 + 23) of them are known to be located in the Small and Large Magellanic Cloud (LMC) respectively. Pulsars in GCs tend to have a more complicated evolutionary history due to the high probability of exchange interactions. Classifying pulsars into different categories is a useful tool for building models to explain the different types of pulsars and their evolutionary history. The two most readily available observables after discovering a pulsar are its spin-period (P) and its spin-period derivative ($\dot{P}$). Therefore, a useful tool is to plot these pulsars in a $P - \dot{P}$ diagram. An example is shown in figure 1.4. As discussed before in section 1.4, we can use P and $\dot{P}$ to get an approximate value for the characteristic age τ_c of the pulsar and also calculate an order of magnitude estimate of the neutron stars surface magnetic field $B_{\rm surf}$. The $P - \dot{P}$ diagram is analogous to the Hertzsprung-Russel diagram which tracks the evolutionary history of conventional stars by plotting the luminosity or absolute magnitude of stars as a function of its effective temperature or color. From figure 1.4, it is clear that there are two islands of pulsar distributions. Pulsars on the right side of the figure are called normal pulsars or slow pulsars. They tend to have a high value for spin-period between

---

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

[5] See `http://www.naic.edu/~pfreire/GCpsr.html` for an updated list.

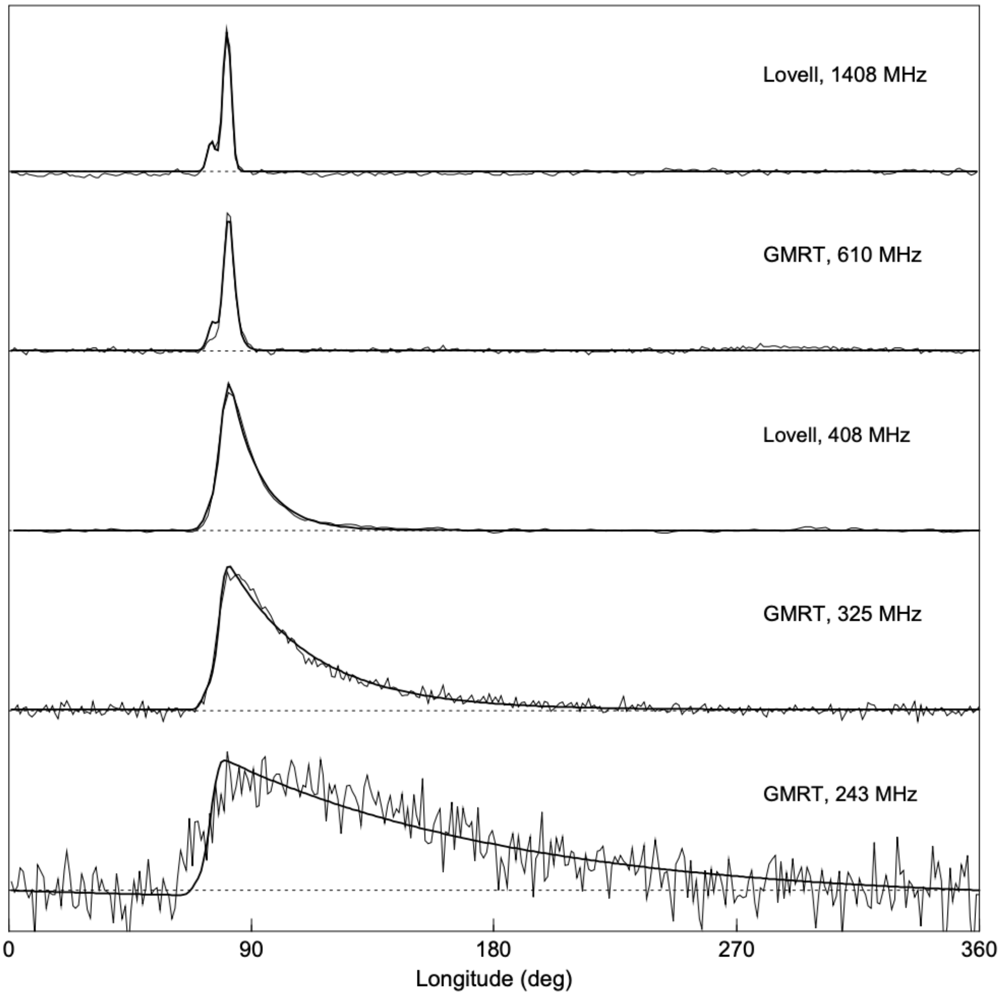


Figure 1.3: An example demonstrating the effects of interstellar scattering on multi-frequency observations of PSR B1831-03. Solid lines show exponential fits to the pulse profile. It is clear that the effects of scatter broadening is more prominent at lower frequencies. Figure taken from Löhmer et al. (2004).

0.1 and a few seconds, high $\dot{P} \sim 10^{-15}$ and a magnetic field of 10^{11-13}G. Pulsars located in the bottom left are called millisecond pulsars (MSPs) as they spin much faster with spin-period values in the order of milliseconds. They also tend to have a lower values of $\dot{P}$ and inferred magnetic field strengths. In the following sections, we will briefly describe the different types of pulsars currently known.

1.6.1 Young Pulsars

The evolutionary history of each pulsar is expected to be slightly different. However, we present a general picture which is consistent with current observations. Pulsars are expected to be born in the top left side of the diagram with a high birth spin-period of 20 ms, high $\dot{P} > 10^{-15}$ and strong inferred magnetic fields $B_{\text{surf}} \sim 10^{12-13}$G (see for e.g. Johnston & Karastergiou 2017). These pulsars tend to have a low characteristic ages τ_c in the range of 1 - 100 kyr. An important evidence confirming their young age is the high number of supernova remnant (SNR) associations in this part of the diagram (see for e.g. Staelin & Reifenstein 1968; Camilo et al. 2002). Supernova remnants are only expected to last around 10^{4-5} years. An additional benefit of associating a pulsar with a supernova remanant is it is possible to get multiple independent measures of age of the pulsar (see for e.g. Xu et al. 2005). A classic example of a pulsar in this category is the Crab Pulsar (Staelin & Reifenstein, 1968) found in the Crab Nebula, a supernova remnant from the supernova SN 1054 which is only 967 years old.

1.6.2 Normal Pulsars

As discussed earlier in section 1.4, pulsars tend to spin-down over time as they lose rotational kinetic energy and move to the bottom right side of the diagram in a timescale of around 10^{5-6} years. This is where we see the vast majority of the known pulsars mainly because they spend most of their lifetime in this region. They tend to have spin-period values ranging from 0.1 s to a few seconds. As expected, they also have a larger characteristic ages ranging from 300kyr $\leq \tau_c \leq$ 1Gyr. Their surface magnetic field gets weaker over time and if no external influences are involved then after about 10^7 years, the pulsar reaches the so called 'death line' where the drop in electrostatic potential becomes too steep in order to maintain coherent radio emission (Chen & Ruderman, 1993) and the NS is no longer detectable as a pulsar. The exact location of the pulsar death line is model dependent and a few pulsars are known to lie beyond the death line. An example of such a pulsar is PSR J2144-3933 (Young et al., 1999) posing questions on our model assumptions or the nature of the pulsar emission mechanisms.

1.6.3 Millisecond and Binary Pulsars

In the bottom left hand part of figure 1.4, we see a secondary population of pulsars which spin with a very short spin-period in the order of milliseconds. As mentioned earlier, these are called *millisecond pulsars* or *recycled* pulsars. These can be differentiated from young pulsars as they have much lower values of $\dot{P} < 10^{-17}$. In this work, we adopt the definition of Ng (2014) where an MSP is defined as a pulsar with $P <$ 30ms

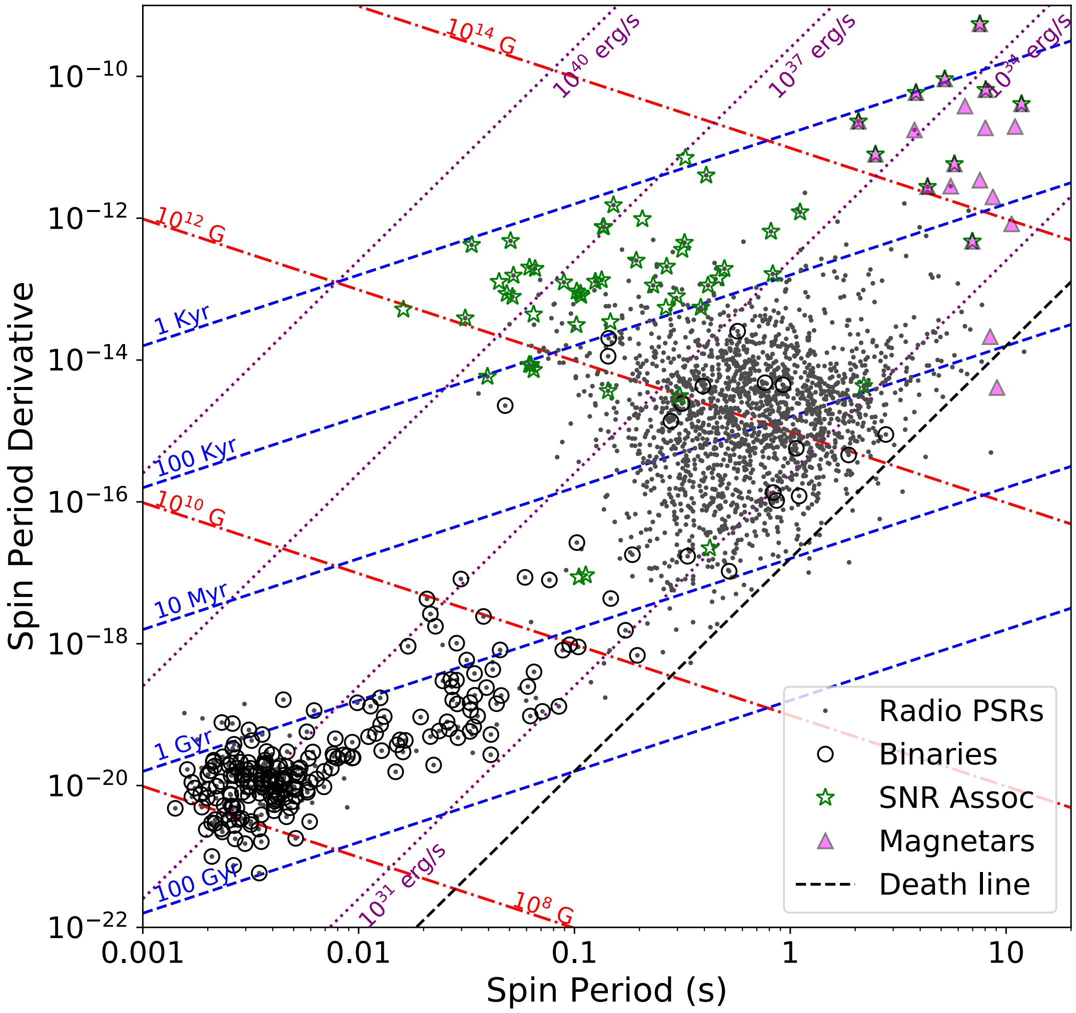


Figure 1.4: A $P - \dot{P}$ diagram of the known pulsars listed in the ATNF Pulsar Catalog (PSRCAT; v1.65; Manchester et al. 2005). We have only included Galactic field radio pulsars (i.e these pulsars not associated with a Globular cluster) Pulsars are plotted as black dots. Points with black circle around them display known binary pulsars. Purple triangles highlight the known magnetars as listed in the McGill Online Magnetar Catalog (Olausen & Kaspi, 2014). Pulsars with known SNR associations are plotted as green stars. Lines of constant spin-down luminosity $\dot{E}$, surface magnetic field B_{surf} and characteristic age τ_c are shown in dotted purple, dashed dot red and dashed blue respectively. In the bottom right region of the diagram we show the pulsar 'death line' defined here based on equation 9 of Chen & Ruderman (1993) which marks the approximate region where pulsars are expected to cease their radio emission.

and $\dot{P} < 10^{-17}$. MSPs also tend to have lower values of inferred magnetic field strength around 10^{8-9} G. Another distinguishing feature is that around 67% of MSPs are found in binary systems compared to 2% of normal pulsars indicating that the evolution of binary pulsars and MSPs should be interlinked. There are multiple formation channels for MSPs. We refer the readers to Alpar et al. (1982); Bhattacharya & van den Heuvel (1991); Phinney & Kulkarni (1994) for a more in-depth review. We present here a simplified version of the standard recycling model. Our evolutionary scenario begins with two Main Sequence (MS) stars bound together in a binary orbit. The more massive star evolves off the MS first and forms a pulsar with the scenario described in section 1.2. If the SN explosion disrupts the binary, then we are left with an isolated young pulsar and a runaway star. If the SN explosion did not disrupt the binary then we are left with a young pulsar in an orbit around a MS star. The pulsar then proceeds to spin-down over time as a normal pulsar over the next 10^{6-7} years. Eventually, the secondary star reaches the end of its MS life and expands to form a red giant. During the red giant phase, the star starts filling its Roche-lobe. Any matter that passes beyond its Roche-lobe gets transferred to the companion via a process known as Roche-lobe overflow (RLO). This process transfers matter and angular momentum towards the pulsar thereby increasing its spin-period. This process is also called ‘recycling’. During the mass transfer, an accretion disk is formed around the pulsar. Frictional heating from the in-falling matter produces thermal X-ray emission and during this stage the system can be detected as low, intermediate or high mass X-ray binary depending on the initial mass of the progenitor star (Tauris & van den Heuvel, 2006). This recycling process is also believed to be the reason for the reduced magnetic field strengths of MSPs (Shibazaki et al., 1989) and the high degree of circularisation of the orbit (Phinney & Kulkarni, 1994). If the companion’s mass is sufficiently low then once the outer layers of the star are shed off in a mass transfer phase that lasts around 10^8 years, then we are left with a WD in orbit around a MSP. Low mass companion systems with $m_c < 0.5 M_\odot$ typically contain Helium white dwarfs (He-WD). These systems also tend to be the most circularised with eccentricities ranging between $10^{-5} \leq e \leq 10^{-1}$. Intermediate mass binary systems tend to have carbon-oxygen (CO) WD or oxygen-neon-magnesium (ONeMg) WD (e.g. Tauris et al. 2011). Given that, MSPs tend to go through a long recycling phase these systems tend to be older with characteristic ages of a few Gyr. These have been independently confirmed using measurements of WD cooling ages (van Kerkwijk, 1996).

An alternate possibility is if the mass of the companion is high enough to eventually undergo its own SN explosion and if the binary orbit survives, then this leads to the formation of a rare Double Neutron Star system (DNS). The second supernova explosion tends to provide an asymmetric kick which either disrupts the binary or if the binary survives, it decircularises the orbit. Therefore such systems are typically found with high eccentricity values (Ihm et al., 2006). So far there are only around 20 known DNS systems [6]. DNS systems have been invaluable for testing gravity in the strong-field regime. Some famous examples include the discovery of the first DNS PSR B1913+16

---

[6] `https://www.atnf.csiro.au/research/pulsar/psrcat/`

(Hulse & Taylor, 1975b). The measurement of the orbital period decay of this system provided the first indirect evidence of gravitational waves and the authors of this work were awarded the 1993 Nobel Prize in Physics. Another interesting example is the discovery of PSR J0737-3039A/B, the only known Double Pulsar system (Burgay et al., 2003; Lyne et al., 2004), which have provided even more stringent tests on General Relativity (GR) and alternative theories of gravity (Kramer et al., 2006; Breton et al., 2008). An illustration of these different formation channels is shown in figure 1.5.

1.6.4 Nulling, intermittent and mode changing pulsars

For most pulsars during the time scales of our observations we expect to see a pulse for every rotation of the star. However, there are number of sub-classes of pulsars where this is not always the case. Backer (1970b) noticed that the radio emission of some pulsars tend to abruptly switch between 'ON' and 'OFF' states. During an 'OFF' state, no emission is detected from the pulsar. This phenomenon is called *nulling*. The duration of nulling varies for each pulsar and is usually expressed in the form of a quantity called nulling fraction i.e the fraction of pulses not seen divided by the total number of expected pulses. Over the years, several pulsars have been reported to exhibit the nulling phenomenon (Backer, 1970b; Ritchings, 1976; Biggs, 1992; Manchester et al., 2001; Wang et al., 2007). Besides the nulling fraction, the duration of the 'OFF' state also varies a great deal. For example, pulsars also show lack of emission on a much longer timescale from hours to days or even years in some cases and these are called intermittent pulsars (Kramer et al., 2006; Camilo et al., 2012; Lorimer et al., 2012; Lyne et al., 2017). This is in contrast to nulling pulsars which tend to switch off for a few pulse periods or minutes. Intermittency has also been linked to spin-down of a pulsar. The difference in the spin-down rates between the two states is pulsar dependent. For example, Kramer et al. (2006) reported an increase in the spin-down rate of PSR B1931+24 by a factor of ~ 1.5 when the pulsar is in its 'ON' state as compared to its 'OFF' state, whereas Camilo et al. (2012) reported a factor of ~ 2.5 increase for PSR J1841-0500. A closely related phenomenon is called *mode changing*. This was first noticed by Backer (1970a) in PSR B1237+25 where the emission process transforms between two or more discrete states which is seen in the form of changes to the integrated pulse profile of the pulsar. The emission mechanism usually favors one state typically called *normal* state for most of the time and can sporadically change to an *abnormal* state at other times from times ranging from minutes to hours. Mode changes have also been reported to be correlated with changes in spin-down rate of the pulsar (Lyne et al., 2010). It is important to note that mode changing, nulling and intermittency are expected to be closely related to the pulsar emission mechanism process which is yet to be fully understood and is unlikely to just be caused by propagation effects discussed in section 1.5.

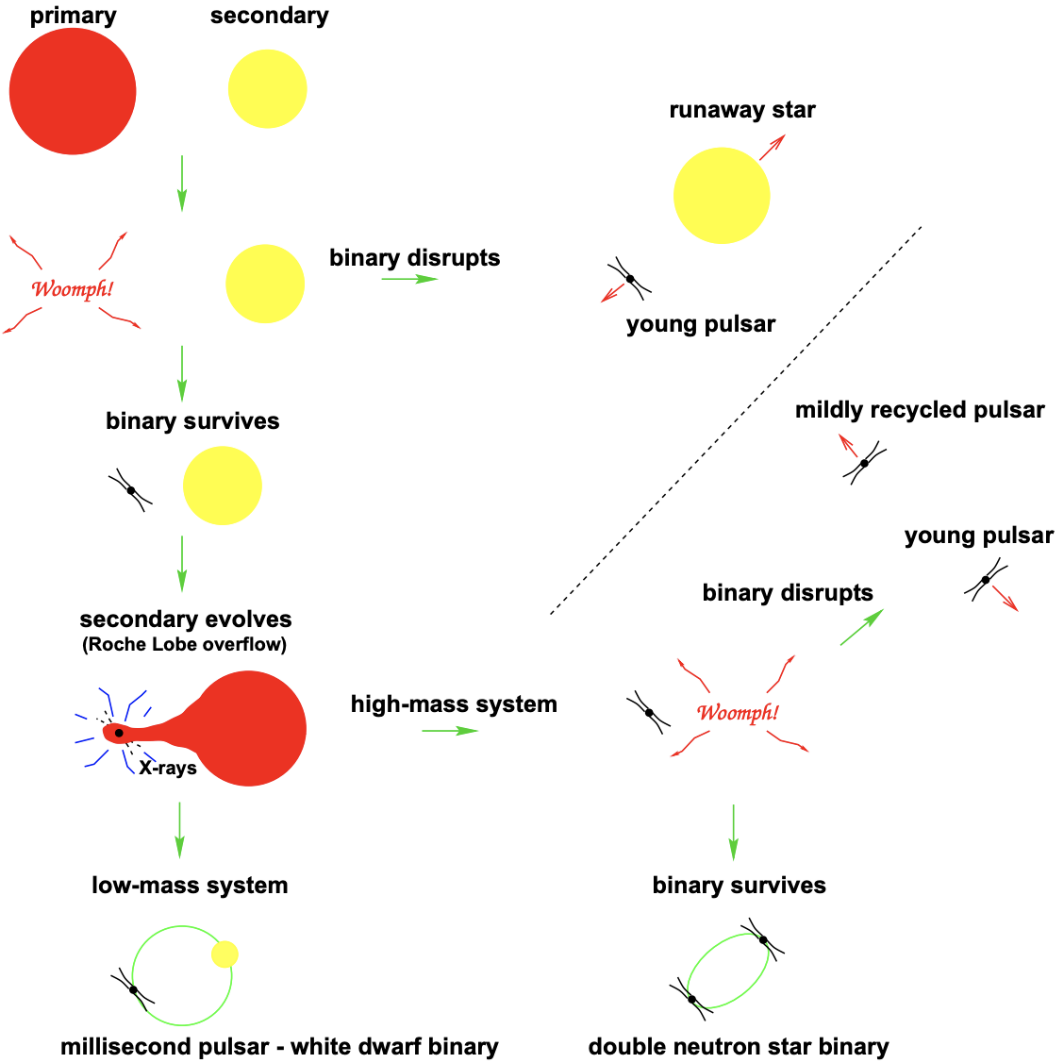


Figure 1.5: An illustration showing the different formation channels for creating an MSP. Figure taken from Lorimer (2008).

1.6.5 Magnetars

Found in the top right side of the $P - \dot{P}$ diagram, magnetars are isolated neutron stars which possess a high magnetic field strength of $\mathrm{B_{surf}} \sim 10^{14-15}$G and relatively long spin-period between 2 and 12 seconds. These systems were initially found in Gamma-Ray and X-ray searches as Soft Gamma Repeaters (SGRs) and Anomalous X-ray Pulsars (AXPs) and were later proposed to be related to highly magnetised pulsars by Duncan & Thompson (1992) when conditions for efficient helical dynamo action are met during the first few seconds of the formation of the NS. The emission mechanism of magnetars is believed to be powered by the decay of their ultra-strong magnetic field (Dall'Osso et al., 2012). There are currently 30 Magnetars known,[7] (Olausen & Kaspi, 2014) of which 16 are SGRs and 14 are AXPs. 5 of these magnetars have also been detected in Radio. (Camilo et al., 2006, 2007; Levin et al., 2010; Eatough et al., 2013; Karuppusamy et al., 2020). We refer the readers to Kaspi & Beloborodov (2017) for a detailed review of the present state of magnetar research.

1.7 Pulsar Timing

Most of the scientific applications of pulsars (discussed in section 1.8) are made possible through a technique called *pulsar timing*. A fully complete review of this topic is beyond the scope of the thesis. However, we will briefly discuss the topics which are relevant for our work. For a more exhaustive review we refer the readers to Chapter 8 of Lorimer & Kramer 2012. The fundamental input of pulsar timing is the time of arrival (TOA) of a signal from a radio pulsar. Since pulse profiles have a certain width, a TOA refers to a fiducial point in the pulse profile. In order to have a stable pulse profile, we average pulses (tens of thousands in the case of MSPs, and hundred to thousand in the case of normal pulsars), together to form an integrated pulse profile. Following this, we create a noise-free analytical template of the pulse-profile which is created by representing the pulse profile as a sum of multiple Gaussian components (Foster et al., 1991; Kramer et al., 1994). Following this, we improve the accuracy of our TOAs by cross-correlating the observed profile with the noise-free template. The goal of pulsar timing is to account for every rotation of the pulsar between observations. Therefore, this requires maintaining an accurate timing standard. Typically TOAs are measured from the local time at the observatory which is maintained by hydrogen maser clocks. However, due to the rotation of the Earth this frame of reference is non-inertial. Therefore, the next step is to transform the TOAs from this topocentric frame to an inertial frame of reference which is typically the Solar System Barycenter (SSB). We then do a least-squares fit analysis in order to minimise the difference between the predicted and observed TOAs. This difference is called timing residual. A good fit will show a Gaussian distribution of timing residuals with mean equals to zero and root mean square that is comparable to TOA uncertainties. Initially, we only fit for period and pulse reference phase which are expected to impact TOA precision in the short

---

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

timescale. However, gradually as the dataset expands, we need to introduce additional parameters in order to fully model the rotation of pulsars. These can be divided into three categories namely *astrometric* parameters like proper motion, position, parallax, *spin parameters* like the spin-period and higher order derivatives of spin-frequency (as discussed earlier in section 1.4) and finally *binary* parameters if the pulsar is orbiting a companion star. These different processes introduce systematic signatures in our post-fit TOA residuals which can then be identified and added into our model. An example of the post-fit TOA residuals for the pulsar PSR B1133+16 is shown in figure 1.6. Panel (a) shows the residuals for a perfect timing model with residuals clustered around zero. Panel (b) shows a parabolic increase in the residuals which is caused by underestimating $\dot{P}$ in this case by 4 %. Panel (c) shows the residuals from an incorrect position of the pulsar in this case the declination of the pulsar is offet by 1 arcminute. This creates a sinusoidal signature with a period of 1 yr. Panel (d) shows the residuals for a pulsar when pulsar's proper motion is not accounted in this case $\mu_{\mathrm{T}} = 380\,\mathrm{mas\,yr^{-1}}$. Typically, a pulsar needs to be observed for at-least one year with regular monthly cadences in order to get an unambiguous measurement of the position of the pulsar. This helps in breaking the degeneracy between $\dot{P}$ created by an incorrect position and the intrinsic $\dot{P}$. The 1-yr constraint comes from the orbit of the Earth around the Sun. Once this is done the pulsar is considered as solved and we have a phase-connected timing solution. Timing beyond the initial one-year can continue for special cases for example MSPs are particularly useful for pulsar timing as they tend to have high rotational stabilities making them useful tools for high precision pulsar timing for example in Pulsar Timing Arrays (section 1.8.5). Additionally, relativistic binary pulsars also require long-term timing in order to get higher significance measurements of Post-Keplerian (PK) Parameters (section 1.8.1).

1.7.1 Solving Binary Pulsars

The steps mentioned above assume that the pulsar is isolated. However, if the pulsar is in a binary orbit around a companion, then this introduces an additional Römer delay which needs to be accounted for by introducing additional parameters into our timing model. We usually start by assuming a Keplerian Model for the binary pulsar which introduces five orbital parameters namely the orbital period (P_b), projected semi-major axis ($x = a_p \sin i$), epoch of passage of periastron (T_0), orbital eccentricity (e) and longitude of periastron (ω). The standard timing model used for implementing this is the Blandford & Teukolsky (BT, Blandford & Teukolsky, 1976). However, the BT model has its limitations as for pulsars with low eccentricity values e, ω, T_0 are not well defined, therefore typically an alternate parameterisation called the ELL1 model is used (Lange et al., 2001). The standard technique used to find the orbit of the pulsar is by fitting the orbital parameters of the system to a series of spin-period measurements closed spaced in time in order to sample the different orbital phases of the binary. A software implementation of this is the tool FITORBIT[8] which was used to solve the binary pulsars reported in section 3.3.2. Freire et al. (2001) introduced an

---

[8] `https://github.com/vivekvenkris/fitorbit`

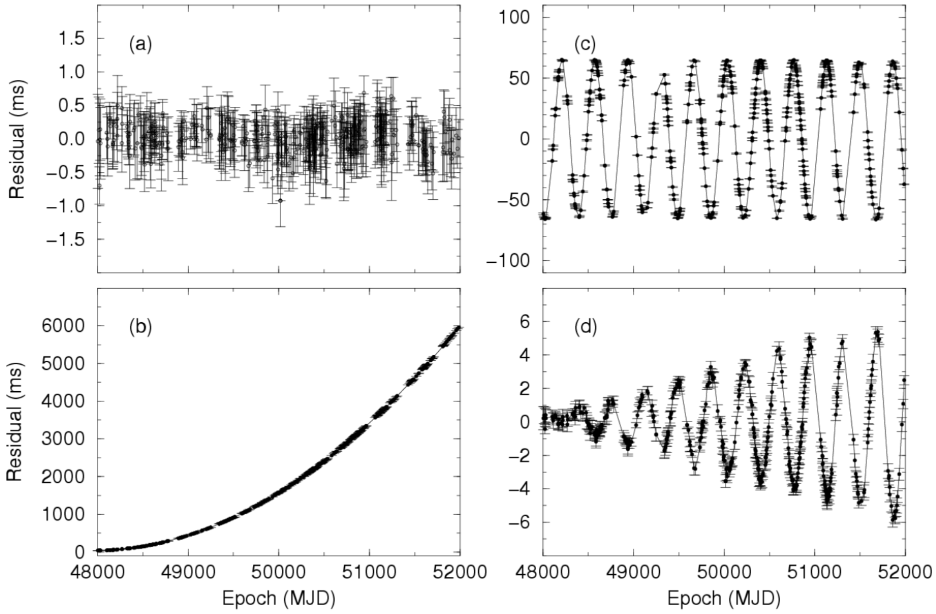


Figure 1.6: Example of Post-Fit TOA residuals for the pulsar PSR B1133+16. See main text for discussion related to each sub-panel. Image taken from Lorimer & Kramer (2012).

alternative technique that can be applied for pulsars which have only a sparse number of detections (for e.g. due to scintillation, see Camilo et al. 2000). This is done by plotting the acceleration of the pulsar versus the spin-period in a period-acceleration plane. Pulsars which have a circular orbit form ellipses in this plane and by finding the best-fit ellipse for our measurements, we can use the centre of the ellipse, the length of the semi-major and semi-minor axis in order to calculate the orbital parameters of the binary. For pulsars in relativistic orbits, a Keplerian model is insufficient to describe the orbit in long-time scales. Therefore, additional parameters called PK parameters are introduced which take into account GR effects. These have been expanded upon in section 1.8.1.

1.8 Scientific Application of Pulsars

In this section, I will describe some of the most interesting scientific applications of pulsars. This list is not intended to be exhaustive but is meant to highlight the diverse range of scientific output that pulsar discoveries can provide. Many of these applications are the primary motivation for much of the pulsar search work done during this thesis.

1.8.1 Strong-Field Tests of Gravity

General relativity (GR) first proposed in 1915 (Einstein, 1915a) is our current best description about the nature of gravity. Over the past century, GR has undergone a wide range of experimental tests. Some examples include the first three classical tests of GR proposed by Einstein (Einstein, 1916) which are presented here in the order in which they were confirmed. The first test to be confirmed was the Perihelion precession of planetary orbits. Einstein (1915b) showed that GR could account for the perihelion advance of Mercury's orbit which could not be explained with Newton's theory of Gravity. The second test to be confirmed was gravitational light deflection. Light travelling close to a massive gravitational source like a star is bent due to the curvature of spacetime around the star. This was confirmed from data taken during the total solar eclipse on May 29th 1919 (Dyson et al., 1920). The last classical test to be confirmed was gravitational redshift. As a consequence of the equivalence principle, light coming out of a gravity well is expected to be redshifted. This has been experimentally confirmed in laboratories (Pound & Rebka, 1959) and in astronomical observations by measuring the redshift in spectral line profiles of Sirius B (Greenstein et al., 1971). Shapiro (1964) proposed a "fourth" test of GR which is called gravitational time delay or Shapiro delay. This is a phenomenon in which light travelling in a gravitational field is delayed compared to a signal which passes through no gravitational field. This was confirmed by Shapiro et al. (1971) by reflecting radar signals from Venus. The idea here was that when Earth, Venus and the Sun were favorably aligned, there would be a measurable extra time delay caused due to the gravitational potential of the sun. These values agreed well with GR upto the 5% limit. Recent measurements of the frequency shift of radio photons to and from the Cassini spacecraft as they passed near the Sun has led to a confirmation of GR upto the 10^{-5} limit (Bertotti et al., 2003). Besides, these four, GR has passed a number of other experimental tests. We refer the readers to Wex (2014) and references within for a more detailed review. The solar-system tests mentioned above tests GR in the quasi-stationary weak-field regime (Wex, 2014) i.e $v \ll c$ and spacetime is close to Minkowski spacetime i.e only very weakly curved. Binary Pulsars on the other hand help us test GR in the quasi-stationary strong-field regime i.e. in addition to non-relativistic speeds one or multiple bodies are strongly self-gravitating which introduces significant spacetime curvature. By using the technique of pulsar timing **(section 1.7)**, we can measure various PK parameters which are phenomenological additions to the standard Keplerian parameters and these measurements can be used to do multiple theory independent tests of gravity. The first successful demonstration of this application was the discovery of the first binary pulsar PSR B1913+16 (Hulse & Taylor, 1975b) which is a double neutron star system in an orbital period of 7.75 hours with an eccentricity of e = 0.62. Within a year, the rate of advance of periastron (PK Parameter) of this system was measured to be $\dot{\omega} = 4.22(4)^{\circ}\mathrm{yr}^{-1}$ (Taylor et al., 1976) which is 35,000 times larger than the GR contribution to the precession of Mercury's orbit $42.98(4)''$ per century (Will, 2014). Additionally, after a few years Taylor et al. (1979) measured two more PK parameters the Einstein Delay amplitude γ_E and the orbital period decay $\dot{\mathrm{P}}_{\mathrm{b}}$. The measurement of the orbital period decay of this system provided the first indirect evidence of gravitational waves as the orbital shrinkage predicted from GR due to the emission of gravitational waves matched exactly with observations (Taylor & Weisberg, 1982). An example of this is shown in figure 1.7 using updated timing observations of PSR B1913+16 lasting over 30 years (Weisberg

& Huang, 2016). In terms of gravitational physics, a more interesting example was the discovery of PSR J0737-3039A/B, the only known Double Pulsar system (Burgay et al., 2003; Lyne et al., 2004). As the name suggests, this systems contains two radio pulsars. Pulsar A is a recycled pulsar with spin-period of 23ms and Pulsar B is an unrecycled slow pulsar with a spin-period of 2.8s. These two pulsars move around in a much more compact orbit of 2.45 hrs with a mild eccentricity of e $\sim$ 0.08. This system is much more relativistic than PSR B1913+16 with $\dot{\omega}$ of nearly $17°\mathrm{yr}^{-1}$. The nearly edge-on inclination of the orbit, has helped in the measurement of three additional PK parameters, the Shapiro range r_s and Shapiro shape, s_s and relativistic spin precession which is measured by using the eclipses of pulsar A at its superior conjunction caused by the magnetosphere of pulsar B. Additionally by timing pulsars simultaneously, we can measure the projected semi-major axis (x) of the orbit of both pulsars which gives the mass ratio R of the system. Therefore, a total of 6 PK parameters were measurable in the case of the double pulsar. Using two PK parameters, we can uniquely determine the mass of the pulsar and the companion by assuming a theory of gravity. These measurements are usually displayed in a so called mass-mass diagram. An example of such diagram for PSR J0737-3039A/B showing all the PK parameters described above is in figure 1.8. Since, we are trying to measure two independent quantities, namely the mass of the pulsar and the companion, N curves in the mass-mass plane, leads to $(N-2)$ tests of a theory of gravity. For a viable theory of gravity, all these PK parameters should meet at a single point within their uncertainty ranges. The double pulsar was used to do 4 independent tests of GR which it successfully passed within a 5% uncertainty range (Kramer et al., 2006). Additionally, Breton et al. (2008) performed a fifth test of GR by measuring the relativistic spin precession rate of pulsar B which was found to be consistent with GR at the 13% uncertainty range.

Besides gravity tests using DNS systems, NS-WD systems are useful for testing the Strong Equivalence Principle (SEP) (Damour & Schaefer, 1991; Stairs et al., 2005; Archibald et al., 2018; Voisin et al., 2020) which is a prediction of GR but is expected to be violated in multiple alternative theories of gravity. Additionally such systems can also be used to put constrains on *dipolar* gravitational waves which are predicted by alternate theories of gravity like the scalar-tensor family of gravitational theories (Damour & Esposito-Farese, 1992; Damour & Nordtvedt, 1993). *Dipolar* gravitational waves if they exist will be more noticeable for NS-WD or NS-BH binaries due to the large difference in their gravitational self energy. Some of the current best constrains of scalar tensor theories of gravity using pulsars are from observations of PSR J1738+0333 (Freire et al., 2012), PSR J0348+0432 (Antoniadis et al., 2013) and the pulsar triple system PSR J0337+1715 (Archibald et al., 2018; Voisin et al., 2020).

1.8.2 Probing the Galactic pulsar population

Despite the discovery of over 3100 pulsars (Manchester et al., 2005), plenty of pulsars still await discovery. Faucher-Giguère & Kaspi (2006) estimated that around 10^5 active radio pulsars should be beaming towards Earth. An effective way to sample them is to do an all-sky pulsar survey. Despite best intentions, all pulsar surveys suffer

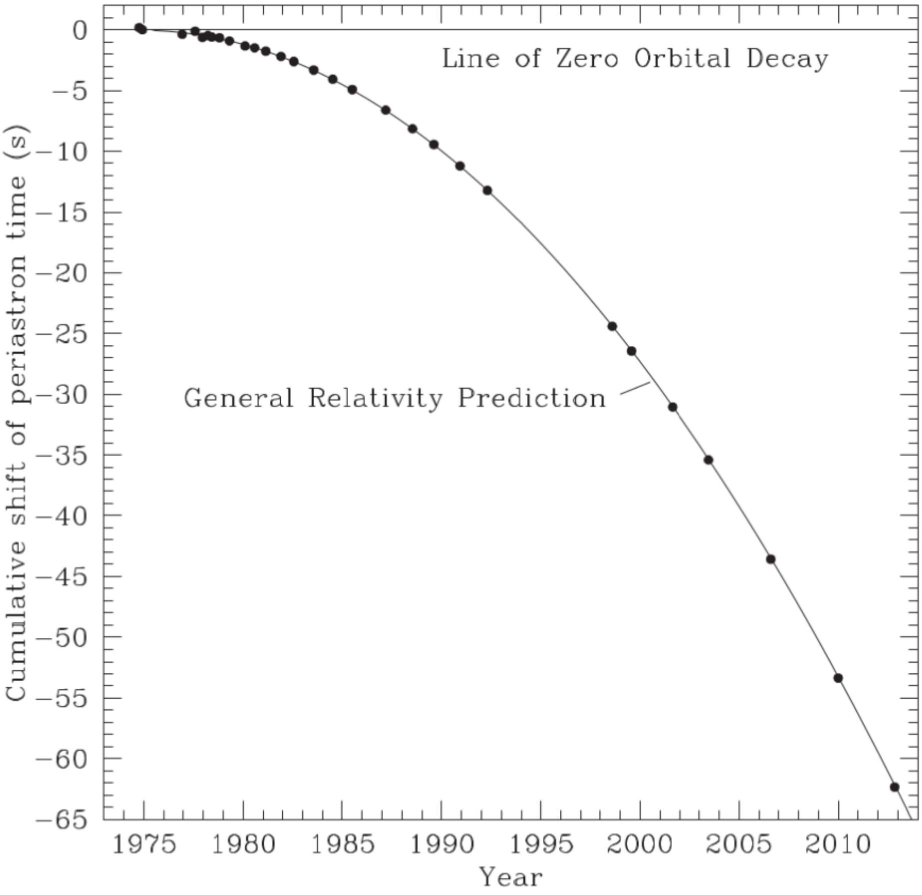


Figure 1.7: The cumulative shift in periastron passage of PSR B1913+16 as a function of time. The horizontal line represents the line of zero orbital decay assuming Newton's theory of Gravity whereas the GR prediction assumes a shrinkage of orbit due to the emission of quadrupolar Gravitational waves. Data points shown in black match with error bars too small to be visible match precisely with the prediction from GR. Image taken from Weisberg & Huang (2016).

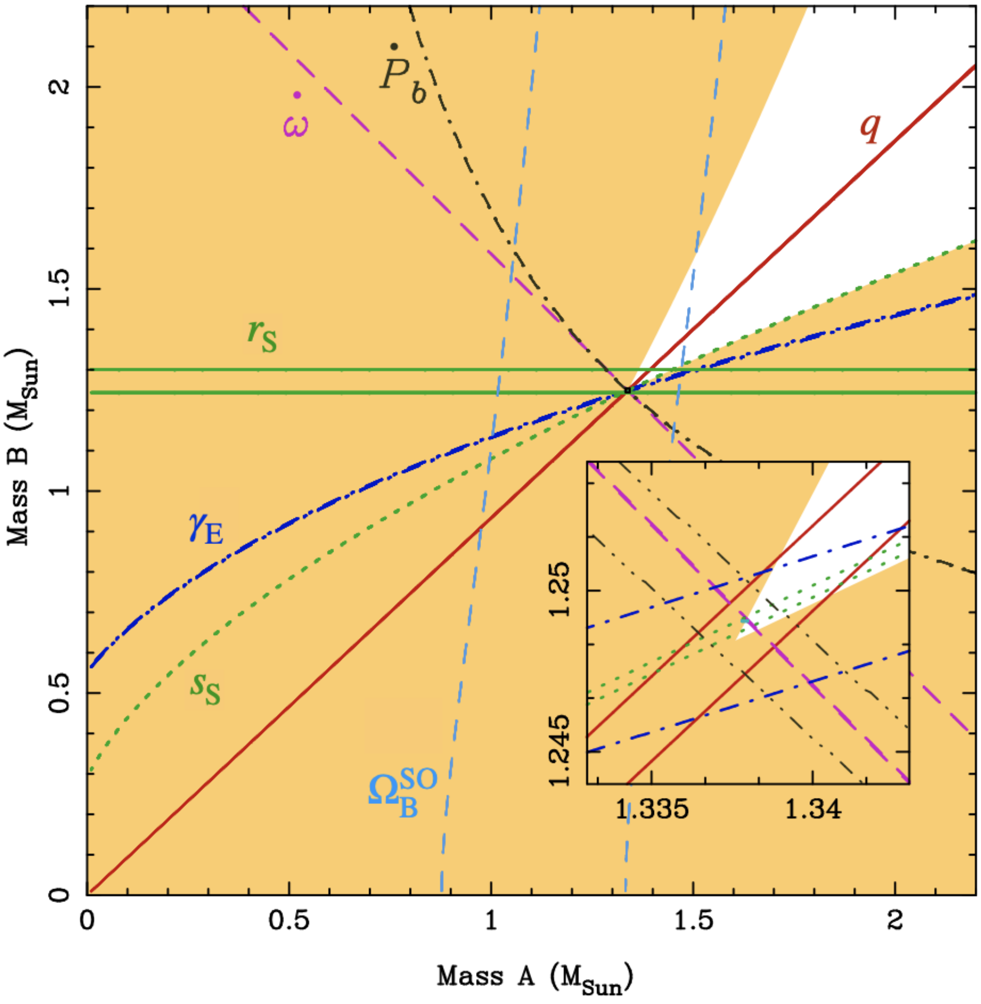


Figure 1.8: Mass-Mass diagram of PSR J0737-3039A/B assuming GR. The shaded regions are excluded by Keplerian mass functions of the two pulsars as $\sin i \leq 1$. The inset shows the overlap between different PK parameters which gives a common solution for the masses of the pulsar and companion. Figure courtesy of Michael Kramer.

from selection effects. For example, from the modified radiometer equation given in equation 1.14, we know that the minimum detectable flux density of a pulsar survey is dependent on the integration time of the observation, gain of the telescope, frequency bandwidth and the system temperature among others which sets a luminosity and distance threshold for the detectable pulsar population:

$$S_{min} = \frac{(\mathrm{SNR}_{\mathrm{min}})\,\beta T_{\mathrm{sys}}}{G\sqrt{n_{\mathrm{p}} t_{\mathrm{obs}} \Delta_f}} \sqrt{\frac{\delta}{1-\delta}} \ \mathrm{mJy}, \tag{1.14}$$

where β is the digisation degradation factor, T_{sys} is the system temperature in Kelvin, G is the Gain of the telescope in KJy^{-1}, n_{p} is the number of summed polarisations, Δ_f is the frequency bandwidth, t_{obs} is the observing time, δ is the duty cycle of the pulsar. SNR$_{\mathrm{min}}$ is the minimum signal to noise ratio threshold which is given as:

$$\mathrm{SNR}_{\mathrm{min}} = \frac{\sqrt{\ln[\mathrm{n}_{\mathrm{trials}]}} - \sqrt{\pi/4}}{\sqrt{1-\pi/4}}, \tag{1.15}$$

where n$_{\mathrm{trials}}$ is the number of independent trials in the pulsar search pipeline.

Additionally, an improved frequency resolution of the survey can help us correct for dedispersive effects better which improves our sensitivity to the detectable MSP population. Long integration time observations while increasing our sensitivity to slow pulsars tends to adversely effect our ability to detect relativistic binary pulsars in compact orbits due to the apparent change in the spin-frequency of the pulsar in an observation caused by Doppler modulation as the pulsar rotates around its orbit. In chapter 5, we will discuss an algorithm that compensates for this effect at the cost of higher computation time. Periodicity searches using the Fast Fourier Transform (FFT) algorithm can hamper our abilities to detect slow-spinning pulsars with a narrow duty cycle (Lazarus et al., 2015; Morello et al., 2020). Only by trying to account for these selection effects by building bigger telescopes, investing in better electronics and improving our search algorithms or by incorporating these selection effects into our models can we get a better understanding of the population of the Galactic pulsar population. Much of the work in this thesis involves the reprocessing of observations from the galactic plane component (HTRU-S Lowlat) of the High-Time Resolution Universe Pulsar (HTRU) survey (Keith et al., 2010). We discuss the implication of our results on the detectable fraction of PSR-BH binaries in the Galactic-plane sky region observed by LOWLAT in sections 3.3.8 and 3.3.9 respectively.

1.8.3 Probing the large-scale structure of our Galaxy and ISM

As described earlier in section 1.5, pulsar signals suffer from various propagation effects like dispersion, scattering and scintillation as it travels through the ISM. Dispersion and scattering directly sample the line-of sight electron distribution. Therefore pulsars with independent distance measurements from parallax, or hydrogen absorption measurements, or from an association with an SNR can be used to calibrate an electron density model like the one discussed in section 1.5.1. Pulsar signals are also highly polarised. The magnetic field of the galaxy acts as a Faraday screen which rotates the position angle of the linear polarisation by an amount proportional to the the electron content times the line of sight component of the Galactic field. By measuring the Rotation measure (RM) of a wide range of pulsars, we can use it to probe the large-scale magnetic structure of the galaxy (e.g. Han 2013).

1.8.4 Matter at ultra-high densities

Pulsars can be used to probe the behaviour of matter at ultra-high densities by measuring it's equation of state. A number of equation of state models have been proposed (see Lattimer 2012 for a review). These models give direct observational predictions of the masses and radius of neutron stars which can then be tested by measuring the mass and radius of the pulsar through the techniques described earlier in section 1.2. An example of such a system is PSR J0348+0423 with a measured mass of $2.01\pm0.04\mathrm{M}_{\odot}$ (Antoniadis et al., 2013) which has already ruled out several equation of state models. Another phenomena of pulsar emission called *glitches* can also be leveraged to probe matter at these densities. Glitches are observed rotational instabilities in the pulsar signal and is likely to be related to the superfluid nature of the liquid neutron star interior. These are usually observed in young pulsars where we see a small sudden increase in the rotation rate followed by an exponential decay back to its pre-glitch rotation rate. This relaxation rate can be thought of as a study of 'neutron-star seismology'. We refer the readers to Baym et al. (1969); Anderson & Itoh (1975); Ruderman (1976); Lyne (1992) for a more in-depth review of the different proposed models for *glitches* and their implications for the interior of a Neutron star.

1.8.5 Direct Detection of nanohertz Gravitational Waves

Measuring the orbital period decay of a relativistic binary pulsar gives us an indirect method to infer the presence of gravitational waves. However, pulsars can also be used to directly detect gravitational waves. This can done by using pulsar-timing arrays (PTAs). MSPs are known to be extremely stable rotators. Therefore, by observing a bunch of MSPs with extremely high timing precision accuracy of about 100 ns and by looking for correlated quadrupole noise signatures in the arrival times of MSPs spaced far apart can help us get a direct detection of Gravitational waves in the nano-Hertz regime (e.g. Jenet et al. 2005). These gravitational waves are believed to be emitted by the merger of super-massive black holes during early galaxy formation (Foster & Backer, 1990; Jaffe & Backer, 2003; Sesana, 2013). PTAs are complementary to ground-based gravitational wave detectors which are sensitive to higher frequency gravitational waves emitted by the merger of stellar-mass black holes (Abbot et al., 2016) and NS-NS mergers (Abbot et al., 2017). Observing a large sample of MSPs requires the coordination of multiple radio telescopes around the world. Three Different PTA projects are currently ongoing. These are the Parkes Pulsar Timing Array (PPTA, Manchester et al. 2013), the European Pulsar Timing Array (EPTA, Kramer & Champion 2013) and the the North American Nanohertz Observatory for Gravitational Waves (NANOGrav, McLaughlin 2013). These projects combined together form the International Pulsar Timing Array (IPTA, Manchester & IPTA 2013). While no detection of gravitational waves have been made yet, the current most stringent limits for nHz gravitational wave amplitudes have been set by Shannon et al. 2015. The ongoing monitoring of PTA pulsars over the next decade especially by new telescopes like MeerKAT[9] and the SKA[10] will be crucial for detecting gravitational waves.

1.9 Thesis Outline

This thesis describes current ongoing efforts to find relativistic binary pulsars in Parkes Observations of the Galactic-Plane Component of the HTRU Pulsar Survey. This includes the development of novel pulsar search algorithms which has increased the sensitivity to compact binaries in long observations. The primary motivation for this was to discover exotic binary pulsars which can help us test GR and alternate theories of gravity in the quasi-stationary strong-field regime. The purpose of the introduction was to set up the context for the primary motivation for this work and to highlight the diverse range of science that new pulsar discoveries can provide. The rest of the thesis is structured as follows.

In **Chapter 2**, we describe the basics of acquiring pulsar data from a single-dish radio telescope and the standard techniques applied in a typical Pulsar-Search Pipeline. We conclude this chapter with an overview of several past and current galactic-plane pulsar surveys.

In **Chapter 3**, we describe a GPU-Based Pipeline of the template-bank algorithm pipeline developed as a part of this PhD which has been built to search coherently for circular orbits in the HTRU-S Lowlat pulsar survey. We applied this algorithm to search for PSR- Stellar Mass BH binaries within our galaxy and to the best of our knowledge this is the most sensitive and widest search for PSR-BH binaries ever conducted in the southern sky. We also present our results which includes the discovery of 20 new pulsars in the re-processing of this survey including a new MSP J1743-24. We additionally also place flux density and detectable fraction limits for orbital parameters of PSR+BH binaries in the galactic plane visible to LOWLAT.

In **Chapter 4**, we describe a novel Semi-Supervised Generative Adversarial Network (SGAN) which is a machine learning algorithm that has been used to classify pulsar candidates. This is the first attempt in literature for applying a Semi-Supervised approach and we believe algorithms like these will be crucial for solving the pulsar candidate classification problem for large datasets produced from telescopes like MeerKAT and eventually SKA.

In **Chapter 5**, we push the abilities of binary search algorithms further by developing the first fully coherent Five-Dimensional Pulsar Search Pipeline which can search coherently for binary pulsars in elliptical orbits. We describe this new approach and compare them to standard acceleration, jerk search and circular binary search pipelines

---

[9] `https://www.sarao.ac.za/gallery/meerkat/`

[10] `https://www.skatelescope.org/`

typically used in literature. Our results show that Elliptical searches for pulsars orbiting Intermediate Mass BH is computationally feasible for observations shorter than 2-hrs in Globular Cluster Observations with an eccentricity limit of $e \leq 0.1$.

Finally, in **Chapter 6**, we conclude our work and detail possible future research paths that can be undertaken from the results of this thesis.

Chapter 2

Pulsar Searching

Contents

2.1 Data acquisition

Pulsars can be detected at radio frequencies with observed flux density at 1.4 GHz ranging from µJy to a few Jy (Manchester et al., 2005). Therefore, in order to detect a vast majority of the pulsars within a reasonable time frame requires a telescope with a large collecting area. There are several radio telescopes currently been used for pulsar searching including the 64-m Parkes Radio Telescope (Australia), the 100-m Effelsberg Radio Telescope (Germany), the 76-m Lovell Telescope at the Jodrell Bank Observatory (JBO; UK), the Giant Metrewave Radio Telescope (GMRT; India), the 100-m Green Bank Telescope (GBT; USA), the 500-m Five-hundred-meter Aperture Spherical radio telescope (FAST; China) and the MeerKAT Telescope (South Africa) among others. Radio telescopes come in a wide range of shapes (parabolic, spherical, cylindrical) and can be used as a stand alone instrument or as a part of an interferometer. All the data analysed as a part of this thesis were recorded at the Parkes radio telescope which has a paraboloid reflector. Therefore, we will briefly discuss the workings of a typical parabolic reflector radio telescope. Radio telescopes do not have an equal sensitivity in all directions. The optics of a parabolic antennae can be thought of as a circular aperture diffraction grating that creates a characteristic beam shape or antenna power pattern consisting of a number of lobes. The main lobe has the overall maximum sensitivity which peaks at the centre (aligned with the telescope boresight, optical axis of the telescope) and decreases as we move away from the centre. An example of a typical antennae power pattern is shown in figure 2.1. An important quantity which is quoted often in this work is the *full width at half maximum* (FWHM), also called *half power beam width* (HPBW) of a telescope beam which is defined as the angular diameter of the region where the response of the beam falls to half its central peak value (-3 dB). The power pattern as described e.g. Rohlfs & Wilson (2000) can be modelled as a Bessel function with FWHM,

$$FWHM = 1.02\frac{\lambda}{\eta D}\text{rad}, \tag{2.1}$$

where λ is the wavelength of the radio signal, η is the efficiency of the telescope and D is the diameter of the telescope. Another important quantity is the raw antennae sensitivity also known as gain G can be calculated as:

$$G = \frac{A}{2k_B} = \frac{\eta\pi D^2}{8k_B}, \tag{2.2}$$

where A is the area of the telescope and k_B is the Boltzmann constant. $G \propto D^2$. Therefore, large telescopes are highly desirable for pulsar observations. An important difference is that unlike radio imaging, for pulsar observations we prefer to have large telescope beam areas in order to survey the sky quickly. Figure 2.2 shows a schematic of the different stages of signal propagation through a radio telescope before the data is digitised and stored for scientific analysis. We will briefly describe each of this in the next section.

2.1.1 Frontend

The term *frontend* usually refers to the initial part of the system that collects the incoming RF signal, amplifies it and later down converts it into a lower frequency. It typically consists of a feed horn, a low noise amplifier, a bandpass filter and a mixer (e.g. Léna et al. 2012). The RF signal collected by the dish and focused on the receiver is collected by a feed horn which is placed at the focal plane of the telescope where standing waves at certain RF are allowed to form. The electric fields created by the standing waves are then sampled by a probe into two orthogonal polarisations (see Burke & Graham-Smith 2009 for a more in depth review) usually dual linear or dual circular. The signal path from each of the polarisation channels is similar, therefore, in the illustration given in figure 2.2, we show the signal path from only one of them. Since the radio signal is weak, the next step is to amplify the signal using a Low Noise Amplifier (LNA) that has a specific frequency response centered at the

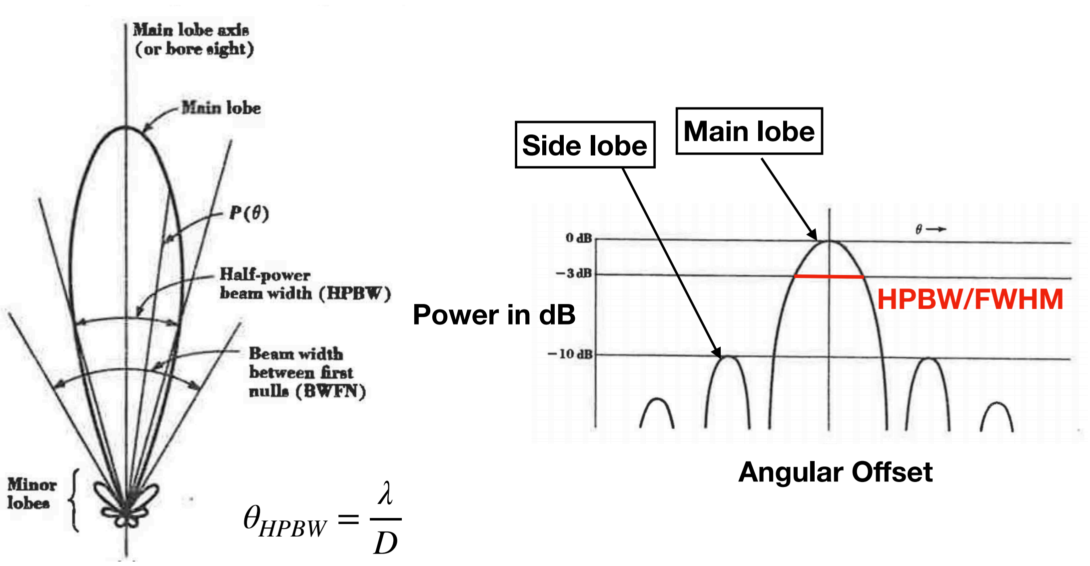


Figure 2.1: Left panel shows the antennae power pattern in polar coordinates plotted in linear scale and the plot on the right side shows the power pattern in cartesian coordinates. The diameter of the main lobe is usuallly referred to as *Full width at half Maximum* (FWHM). This is the angular resolution of the telescope. Figure adapted from Kraus (1966)

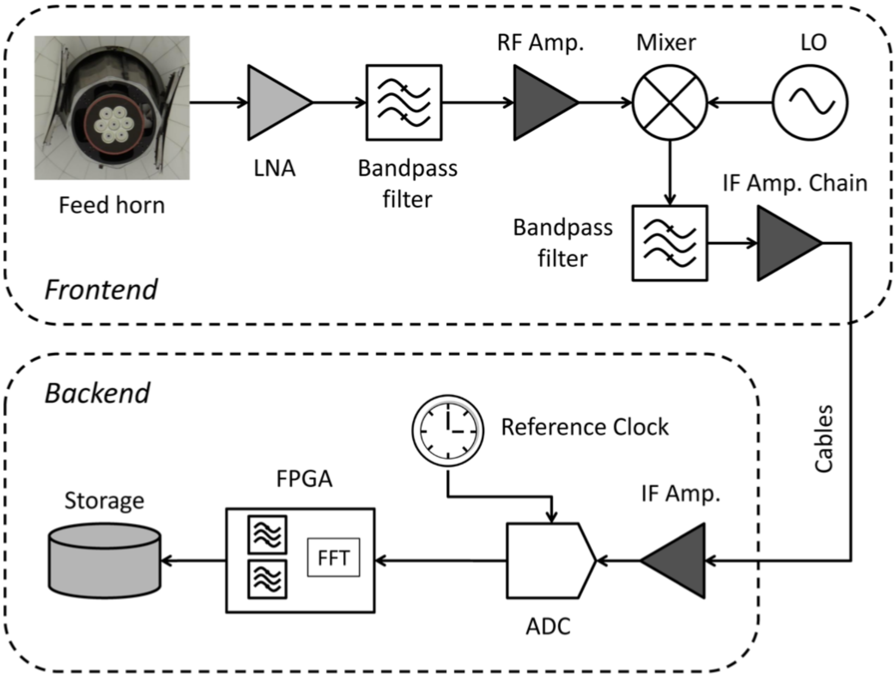


Figure 2.2: Schematic of signal propagation through a radio telescope. Figure taken from (Cameron, 2018)

middle of the observing band. These LNAs are cryogenically cooled to temperatures in the order of few tens of Kelvin to minimise thermal noise. The amplified signal is then passed onto a bandpass filter which samples a specific frequency band and can include additional filters to suppresses any known persistent radio frequency inteference signals before being passed onto an additional RF amplifier. The next step is to down-convert the frequency of the signal using a *mixer* that combines incoming RF signal f_{RF} with a monochromatic signal from a local oscillator f_{LO} through a process known as heterodyning. This is primarily done for two reasons as low frequency signals have lower attenuation losses while transmitting through cables and components operating at low-frequencies are more readily available. The resulting signal at intermediate frequencies (IF) = $f_{\mathrm{IF}} = f_{\mathrm{RF}}$ - f_{LO} is then further isolated by a bandpass filter before amplified by a chain of amplifiers which is necessary to maintain a stable amplification. The signal is then transmitted by cables to where the backend is housed.

2.1.2 Backend

The backend is responsible for digitising, processing and storing thee signal for further analysis. Here, we will describe a specific backend that is commonly used in modern pulsar observations - the polyphase filterbank. The incoming IF signal is sampled at the Nyquist frequency so as to preserve information, and digitised using an Analogue to Digital Converter (ADC). Each digital sample is timestamped using a local high-precision hydrogen maser clock. The ADC converts the signal into a series of n-bit numbers with nanosecond sampling which is then passed onto a Field Programmable Gate Array (FPGA). The FPGA processes discrete blocks of data by calculating a FFT to produce a fourier spectrum every few microseconds. Due to the finite resolution of the FFT, this can cause an input signal at a specific frequency to spread its power over adjacent frequency bins in the power spectrum. This phenomenon is called spectral leakage. To minimise this effect, modern-day backends use polyphase filtering techniques (Lyons, 2011) which filters the signal response by weighting overlapping bins with a sinc function pre-FFT. The combination of polyphase filter and FPGA is referred to as polyphase filterbank. The process also involves channelising the data into narrow frequency channels in order to minimise the effect of intra-channel dispersive smearing (section 1.5.1). The resulting fourier spectra can be handled in multiple ways depending on the scientific application of the data. If the goal is to search for new pulsars, then the data is converted to 'search-mode' where the spectra is integrated to give sampling rates of tens of microseconds with the two orthogonal polarisations summed in quadrature as polarisation information is not required for search-mode observations. If the spin-period and the dispersion measure of the pulsar is already known, then the signal can be folded (section 2.2.5) to form a timing archive. This is called recording in 'pulsar-timing' mode. Finally, the data is converted to a common data format and written to disk. The two most common data formats used for pulsar searching is filterbank[1] and PSRFITS (Hotan et al., 2004). The difference between the two formats

---

[1] Refer to section 3 of `http://sigproc.sourceforge.net/sigproc.pdf` for a detailed description of this format

is based on the way the metadata is stored in the header of the file. PSRFITS is based on Flexible Image Transport System (FITS) (Pence et al., 2010). Most of the data processed during this PhD uses the filterbank format. In brief, a filterbank file contains a two dimensional array of frequency channels vs time. The header file can be thought of as a dictionary which contains metadata like sampling time, frequency bandwidth among others as key value pairs. These files can be processed with a wide variety of Pulsar software suits including SIGPROC[2], SIGPYPROC[3] and PRESTO[4].

2.2 Standard Pulsar Search Pipeline

In this section, we briefly describe the different stages of a typical pulsar search pipeline. An overview is shown in figure 2.3. We additionally give details for specific techniques that were used for the results of this thesis.

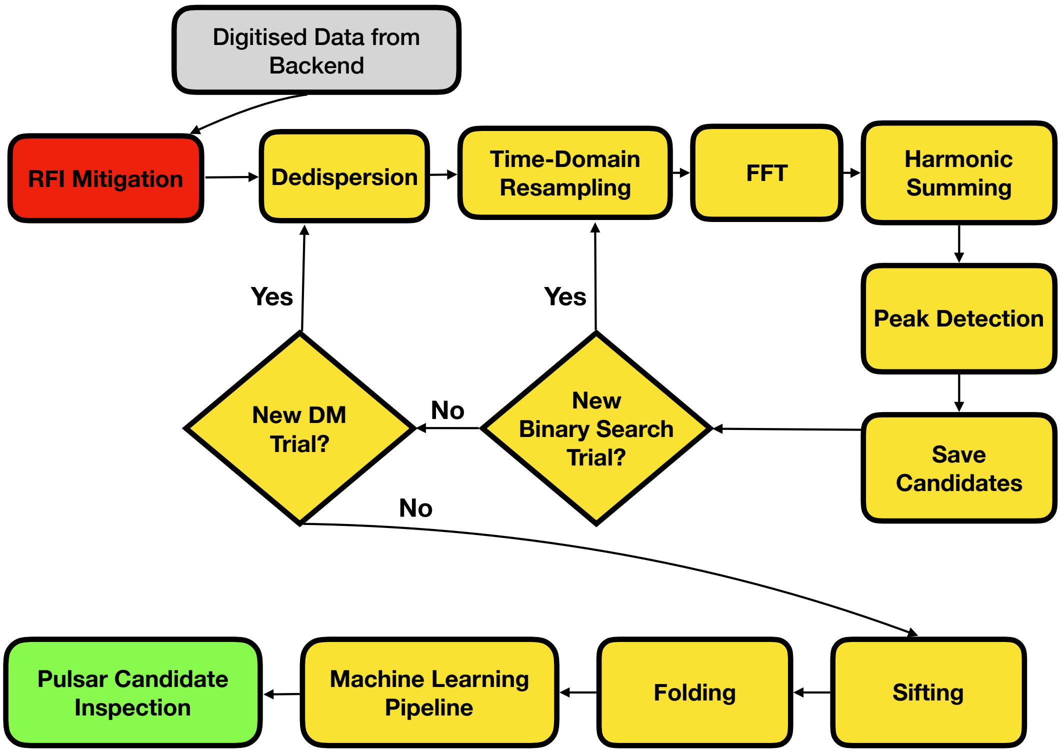


Figure 2.3: Block Diagram showing the different stages of time-domain FFT-Based Pulsar Search Pipeline. See text for more details.

---

[2] https://github.com/SixByNine/sigproc
[3] https://github.com/ewanbarr/sigpyproc
[4] https://github.com/scottransom/presto

2.2.1 Radio Frequency Interference Mitigation

Before we search the data for astrophysical signals, it is important to clean the data. Observations in radio frequencies are adversely affected by various sources like radars, television, AC frequency of voltage mains, mobile networks, satellites, Wi-Fi among others. These are collectively called Radio Frequency Interference (RFI) as they hinder the detection of pulsar signals. Therefore, RFI mitigation is an integral first part of any modern pulsar search pipeline. Various techniques have been described in literature in order to minimise the affects of RFI. We will briefly summarise the most relevant methods which were used in this thesis.

2.2.1.1 Time-Domain

RFI signals are created by terrestrial sources. Therefore, an effective method of removing them is to create a time-series at zero dispersion measure (DM; section 2.2.2) i.e sum all the different frequency channels without applying any delays to form a timeseries. Bursty signals appearing in this time-series are almost certainly caused by signals originating from Earth. These signals can be removed by either clipping the data above a certain threshold or these contaminated signals can be replaced with Gaussian random noise samples. This is an effective method against short-term impulse RFI.

2.2.1.2 Frequency-Domain

Another option for RFI mitigation is to integrate the observation across time to form a bandpass (Power vs Observing Frequency). Similarly, to the time-domain method, data from channels badly affected by RFI can be replaced either with zeros or Gaussian noise samples from the surrounding uncontaminated data. This technique is effective against persistent narrowband RFI.

2.2.1.3 Fourier-Domain

Yet another possibility is to identify RFI signals in the fourier domain. This is done by calculating a FFT of a zero DM timeseries and inspecting the fourier spectrum. Periodic signals that were too weak to be masked by the time-domain and frequency-domain method can now be readily identified and zapped. This is a useful tool to remove higher harmonics of periodic RFI. Genuine pulsar signals are also expected to be periodic in nature, therefore removing periodic RFI at this stage will help in minimising the number of false-positive pulsar candidates (section 2.2.5) at a later stage. These bad frequencies are usually termed as *birdies* in literature and are generally replaced by zeros or white noise in the power spectrum.

2.2.1.4 Multi-beam techniques

If observations are recorded with multi-beam receivers[5], then additional specialised RFI mitigation techniques can be applied which can remove RFI signals down to a lower

---

[5] each receiver points at a separate point in the sky and observations are recorded simultaneously

threshold. The assumption here is that RFI signals caused by terrestrial sources are likely to be visible in most if not all beams of an observation, in contrast pulsar signals are most likely to be seen in a single beam.[6]. An example of a time-domain multi-beam technique which was used during the data processing of this thesis is the 'eigen-vector decomposition' method. Details about this can be found in Kocz et al. (2012). For completeness, we briefly describe the algorithm here. The data from all the beams of the multi-beam receiver are dedispersed (section 1.5.1) at DM= $0\,\mathrm{pc\,cm^{-3}}$. This is followed by calculating the cross-correlation matrix between the different timeseries and calculating its corresponding eigenvalues. By thresholding on the smallest eigen values, we mask signals present in all beams. We applied appropriate thresholds based on how many beams that the signal was detected.[7] This time-domain mask is applied to all beams of an observation where we replace the bad samples with random white noise. We also employed a fourier-domain multi-beam RFI mitigation technique for our data analysis. This involved calculating a FFT on the zero DM timeseries of all beams, and identifying frequencies which are detected in multiple beams. We set appropriate thresholds based on the number of beams that the signal was detected in.

2.2.2 Dedispersion

After RFI mitigation, the next step is to dedisperse the data. As the radio emission from the pulsar passes through the interstellar medium (ISM) before arriving at our telescope receiver, the signal is "dispersed" by the free electron content in the ISM. The amount of dispersion is proportional to the dispersion measure (DM) which is proportional to the integrated column density of free electrons between the pulsar and the observer. This creates a frequency-dependant delay such that lower frequency signals arrive later compared to higher frequency signals. This dispersive delay is calculated from the top of a frequency band to the centre of each frequency channel, and is given as

$$\Delta_{\mathrm{DM,delay_i}} = \left[\frac{4.15\times10^{3}}{\mathrm{t_{samp(s)}}} \times \left(\frac{\mathrm{MHz}}{f_i^2} - \frac{\mathrm{MHz}}{f_{top}^2}\right)\left(\frac{\mathrm{DM}}{\mathrm{pc\,cm^{-3}}}\right)\right]_{\mathrm{nearest-int}}, \tag{2.3}$$

where $\Delta_{\mathrm{DM,delay_i}}$ is the calculated delay in seconds for the frequency channel f_i with respect to the top of the frequency band $\mathrm{f_{top}}$ at a particular DM and for a sampling time $\mathrm{t_{samp}}$ for the observation. We then calculate the nearest integer of this value and slide each channel by the dispersive delay required. After, this we sum across the frequency channels to form a dedispersed timeseries. Mathematically, this can be written as:

$$T_j = \sum_{i=1}^{n_{\mathrm{chans}}} S_{j+\Delta_{\mathrm{DM,delay_i}},i} \tag{2.4}$$

---

[6] There are notable exceptions, for example a bright pulsar like PSR J0835-4510 can be detected in multiple neighbouring beams due to the enhanced sensitivity provided by the beam's sidelobes.

[7] Higher the number of beams in which the signal is seen, lower is our corresponding sigma threshold to mask them.

where T_j is the $\mathrm{j^{th}}$ sample of the dedispersed timeseries. This is called *incoherent dedispersion* as there is still dispersive delay that is unaccounted for within a frequency channel. An example of the dispersive smearing effect is shown in figure 2.4. In the case where the DM of a pulsar is known beforehand, dispersion effects can be completely removed by using a technique called *coherent dedispersion*. We refer the readers to Chapter 5.3 of Lorimer & Kramer (2012) for more details on this technique. Rearranging equation 2.3, we can calculate the dispersive delay within a frequency channel of width Δ_f assuming $\Delta_f \ll f$ where f is the central observing frequency as:

$$\Delta_{\mathrm{chan,delay}} = 8.3 \times 10^3 s \times \mathrm{DM} \times \Delta_f \times f^{-3}, \tag{2.5}$$

In a blind pulsar survey, the true DM of a pulsar is *a priori* unknown, we typically de-disperse the data into multiple trial values. The range of the DM values depend on the part of the sky we are observing with the range provided by an electron density model such as YMW16 (Yao et al., 2017) and the step size is decided based on our acceptable loss of signal to noise ratio (SNR) tolerance level. If the step sizes are too close, computation time is wasted on redundant trials, whereas if the trials are spaced too far apart, then we could miss new pulsar discoveries as the signal would be smeared out. In practise, we use a DM tolerance parameter DM_{tol} that limits the maximum dm smearing time which can be calculated as:

$$\mathrm{DM_{tol}} = (1 + \varepsilon) \geq \frac{W_{\mathrm{tot}}}{W_{\mathrm{eff}}}, \tag{2.6}$$

where $\mathrm{W_{tot}}$ is the total observed pulse width and $\mathrm{W_{eff}}$ is the effective pulse width of pulsar. Values can range from 1.01 (no losses) to 1.35 depending on the use case. We typically use a DM tolerance of 1.11 for most of our analysis in this thesis.

2.2.3 Periodicity Searches

Once, the observation has been cleaned and dedispersed for a given trial DM, the next step is search for periodicities in our data for a possible pulsar signal. There are two standard approaches in the literature to solve this problem. First is by calculating a Fast Fourier Transform (FFT) (Cooley & Tukey, 1965) and running a peak detection algorithm to find the spin-period of the pulsar. The second approach is entirely done in the time-domain and is called Fast-Folding Algorithm (FFA) (Staelin, 1969). The latter is more suited to find slow spinning pulsars with narrow duty cycles. We refer the readers to Morello et al. (2020) for a more in depth review of the FFA algorithm. Here, we will briefly discuss here the workings of a typical FFT pulsar search pipeline which is relevant for the results of this thesis.

2.2.3.1 Discrete Fourier Transform (DFT)

Calculating fourier transforms is one of the most efficient ways to find periodic signals in a dataset. Since a dedispersed timeseries is a collection of non-continuous, independent set of samples we use the discrete fourier transform. For a uniformly spaced timeseries

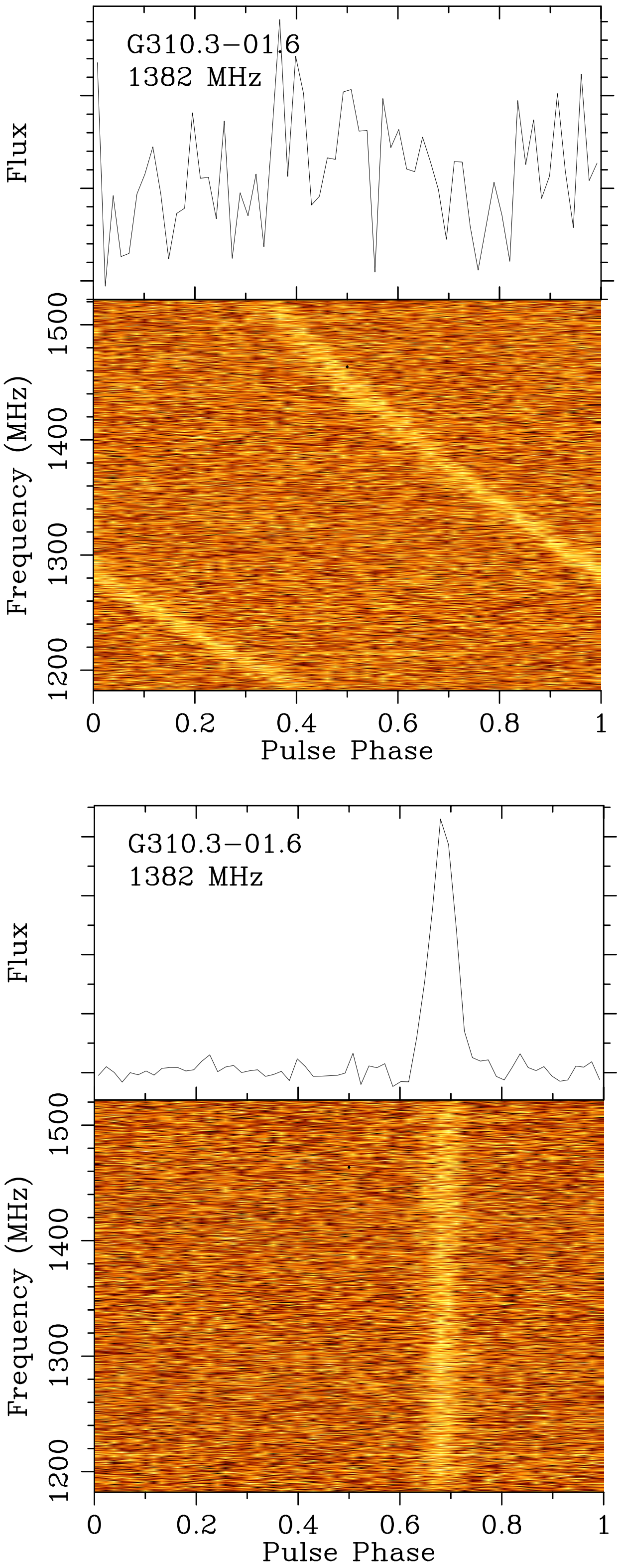


Figure 2.4: An example showing the effect of dispersive smearing on PSR J1357-62. On the top, we show freq-phase and pulse profile plots when no DM corrections have been applied. We can see that signals at the lower end of the frequency spectrum arrive much later compared to higher frequency signals and they follow a quadratic relation. The bottom plot shows the frequency-phase and pulse-profile after we apply corrections for DM (in this case DM = $420.8\,\mathrm{pc\,cm^{-3}}$). This is done by applying appropriate time delays proportional to the DM value to each frequency channel which leads to an improved detection. This process is known as dedispersion.

T$_j$, the kth component of a DFT (see for e.g. Bracewell 2000) can be calculated as:

$$F_k = \sum_{j=0}^{N-1} T_j e^{\frac{-2\pi ijk}{N}}, \tag{2.7}$$

where N is the total number of samples in the timeseries and i $= \sqrt{-1}$, k is usually called the Fourier frequency or wave number. Calculating a DFT for large N is computationally expensive. The above equation is basically a dot product and scales as $\mathcal{O}(\mathrm{N}^2)$. A faster approach is to calculate an FFT (Cooley & Tukey, 1965) which scales as $\mathcal{O}(\mathrm{NlogN})$. A further speed up can be achieved by taking advantage of the fact that the data points in observations are real-valued. Therefore, the fourier transforms are symmetric about the Nyquist frequency ν_{nyquist} such that $\mathrm{F}_{N-k} = F_k^*$ where the latter is the complex conjugate of the former. Therefore, we only need to calculate the amplitudes of half of the fourier frequencies which gains us an additional factor of 2. (Press et al., 1992). The highest frequency that can be detected with an FFT is given by the Nyquist Shannon sampling theorem which is:

$$\nu_{\mathrm{nyquist}} = \frac{1}{2t_{samp}}, \tag{2.8}$$

We then calculate the fourier *power spectrum* by summing the real and imaginary parts as:

$$P_k = \mathrm{Re}(F_k^2) + \mathrm{Im}(F_k^2), \tag{2.9}$$

A limitation of the DFT/FFT algorithm is that due to the finite resolution of the data, the natural frequency of the pulsar need not always lie on the centre of the frequency bin. The response of the FFT in frequency is not uniform. Its maximum for signals falling in the centre of the fourier bin and loses sensitivity as you go towards the edge of a bin. This is called *scalloping* (see for e.g. Middleditch et al. 1993). Ransom et al. (2002) showed that the worst case loss of fourier power could be as high as 41% with an average loss of 23% (van der Klis, 1988). There are multiple options to minimise this. One of the simplest options which we use later in our analysis is to pad the data with the mean of the signal. This is called *mean padding.* Padding adds no power to the data but increases the spectral resolution which reduces the effects of scalloping. This however, does come at the cost of computing a larger FFT. We refer the readers to Ransom et al. (2002) and references within for a more in depth review of this topic and details on other techniques.

All the techniques discussed so far are ideally suited for finding sinusoidal signals in Gaussian white noise. However, pulsars typically have a duty cycle δ ($\delta = W/P_{\mathrm{spin}}$) of a few percent, where W is the width of the pulse. Therefore, the power in the fourier spectrum is distributed between the fundamental frequency and a number of harmonics. To recover the power from these harmonics, a technique called *incoherent harmonic summing* is implemented (Taylor & Huguenin, 1969). It is called incoherent as there is no phase information of the signal. It basically involves stretching the fourier spectrum in powers of two and adding it to the original unstretched spectrum, the idea here is that if the powers in the harmonics are roughly equal then they increase by a factor of 2 whereas the noise in the folded spectrum will only increase by $\sqrt{2}$. We also repeat this procedure to gain power from odd number of harmonics. Since *a priori* in a search, we do not know the duty cycle of a pulsar, we typically do up-to 16 harmonics sums which is sufficient for pulsars with duty cycles $\delta > 0.02$ (Ransom et al., 2002). For wider profile pulsars typically MSPs, adding higher harmonics could reduce sensitivity as it just involves adding noise. Therefore, all the standard pulsar search softwares typically investigate each harmonic fold separately for statistically significant detections before adding them.

2.2.4 Binary Pulsar Search Techniques

Fourier based analysis discussed in the previous sections assume that the spin-period of the pulsar does not change during an observation. Therefore, these are effective for detecting isolated pulsars. However, if the pulsar is in a short period binary orbit, the observed apparent frequency of the pulsar changes due to Doppler modulation. If this effect is unaccounted for, then the signal is smeared onto neighbouring fourier bins which hinders detection. One possibility to correct for this is the *time-domain resampling* technique. The main idea is to add or subtract appropriate delays to the each sample of the timeseries in order to shift the reference frame from the observer to an inertial frame of reference (for e.g. binary system's barycenter). After this, we implement our standard fourier techniques to find the pulsar. Mathematically, this can be written as

$$\tau(t) = \tau_0 \left(1 + \frac{V_1(t)}{c}\right), \tag{2.10}$$

where $\tau(t)$ here represents the time in the barycenter frame, τ_0 is a constant used for normalisation and $V_1(t)$ is the radial velocity of the pulsar and c is the speed of light in vacuum. We have neglected terms in $\frac{v}{c}$ higher than first order. When we are searching for new binary pulsars $V_1(t)$ is a *a priori* unknown, therefore, we need to do a grid search over values of $V_1(t)$ in order to be sensitive to a wide range of binary pulsar orbits. With no prior information, a fully coherent search will require searching over Five Keplerian parameters in addition to searching over period and DM. This is computationally expensive and has never been attempted before. We will expand more on the feasibility of this search technique in chapter 5 of this thesis. A simplified approach to this problem is to first order assume that the line of sight acceleration of the pulsar is a constant during the course of an observation i.e $V_1(t) = a_1 t$. This is called an "acceleration-search" and is the most widely used technique in the literature. For every DM trial, we search over range of acceleration values depending on the orbital parameters of the binary pulsar we want to be sensitive towards. This is usually the most computational demanding part of a modern pulsar search pipeline. Some examples of doing acceleration search using time-domain resampling include analysis of 47 Tucanae Globular Cluster observations (Camilo et al., 2000), the Parkes multibeam pulsar survey (PMPS) (Manchester et al., 2001), re-analysis by Eatough et al. (2013), analysis of the High-Time Resolution Universe (HTRU) survey (Ng et al., 2015; Cameron et al., 2020) among others. An alternative approach is to correct for the binary motion entirely in the fourier domain. This is the so called "correlation-technique" (Ransom, 2001; Ransom et al., 2002) where after calculating a FFT, we use matched filtering in the fourier domain to "sweep up" signal spread over a number of frequency bins into a single bin. An implementation of this technique can be found in the code ACCELSEARCH which is a part of the PRESTO[8] software package. This technique has also widely been used in literature. Some notable examples include its first application which led to the discovery of PSR J1807-2459A, a binary pulsar with an orbital period $\mathrm{P_{orb}}$=1.7 hrs in the Globular Cluster NGC 6544 (Ransom, 2001), eccentric, binary millisecond pulsar PSR J1946+3417 (Barr et al., 2013) and a number of discoveries in the Arecibo Pulsar-ALFA (PALFA) survey (Lazarus et al., 2015). Despite the success of acceleration searches so far, it has its drawbacks. The assumption of constant acceleration is valid only if the observation samples a small fraction of the orbit typically for long orbital period binaries $P_{\mathrm{orb}} \geq 10\, T_{\mathrm{obs}}$ (Ransom et al., 2003; Ng et al., 2015). Therefore, shorter observations are more preferable for finding relativistic binary pulsars. However, the sensitivity of a pulsar scales as $\mathrm{T_{obs}}^{1/2}$, therefore with an acceleration search we have a trade-off between sensitivity and the ability to detect relativistic binary pulsars. Recently, Andersen & Ransom (2018) implemented Fourier domain jerk searches which has increased sensitivity to orbits in the $P_{\mathrm{orb}} \geq$ 5-15 T_{obs} regime. Allen et al. (2013); Knispel et al. (2013) implemented full orbital searches assuming a circular orbit to search for binary pulsars in PALFA and PMPS survey respectively. In chapters 3 and 5, we will compare the sensitivity and use-cases of the search techniques discussed above to the new techniques developed during the analysis done for this thesis. Other less commonly used techniques include "stack/slide" technique (Faulkner et al., 2004) and a phase modulation search (Ransom et al., 2003).

2.2.5 Candidate Sifting and Folding

After searching through the data using the techniques listed above, we are left with a list containing the FFT spin-period, SNR along with the corresponding DM and acceleration value it was detected at. A real pulsar signal can be detected with varying SNR across multiple period, DM and acceleration values. Therefore, it is important in order to save compute time to group together duplicate detections, retaining only the candidate with the highest SNR. This process is known as *sifting*. After sifting, we then fold the observations in phase using the detected parameters to form a folded archive. This is typically called a pulsar candidate. Pulsar candidates are four-dimensional datacubes consisting of time, frequency, rotational phase and power of a signal. We then produce various one-Dimensional and two-Dimensional projections of this datacube in order to visually diagnose if the signal is from a pulsar or not. An example of such a diagnostic plot showing a real pulsar signal discovered from the analysis of this thesis is shown in figure 2.5. The top left hand side of the plot marked as '1' is the pulse profile of the signal. We create this by integrating over both time and frequency axes

---

[8] `https://github.com/scottransom/presto`

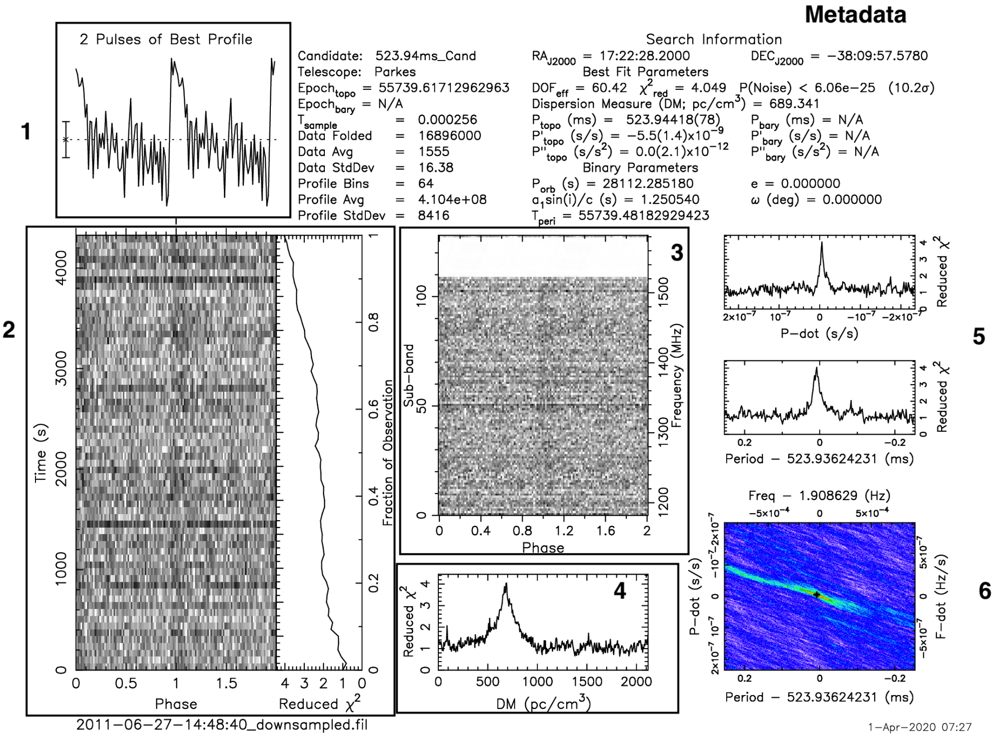
while preserving phase. Pulsar signals can have one or multiple peaks with a narrow or broad duty cycle ranges whereas RFI signals tend to be sinusoidal. The plot marked as '2' is the Phase-Time plot and the plot marked as '3' is the Phase-Frequency plot. Real pulsar signals tend to be persistent (vertical line) across both frequency and time. The plot marked as '4' is the DM curve plot. This is created by dedispersing the data around a few trial DM values from the original DM used to fold the candidate, it then calculates a χ^2 of the dedispersed pulse profile against a horizontal line fit. A large χ^2 indicates a strong deviation. We expect pulsar signals to peak at a non-zero DM value. The plots marked as '5' show χ^2 vs Period and P-dot. This can be thought of as a fit for P and P-dot and we expect real signals to peak at a particular value of P and P-dot. The plot marked as '6' shows the signal in P-Pdot plane. Traditionally, plots like these were visually analysed in order to classify between Pulsar and RFI signals. However, modern pulsar surveys like HTRU-S Lowlat produce tens of millions of candidates in one processing run making it practically infeasible to inspect all of them. Chapter 4 of this PhD thesis presents a novel Semi-Supervised Machine learning algorithm that automates this process.

2.2.6 An overview of Galactic-Plane Surveys

We conclude this chapter with a brief overview of the pulsar surveys conducted so far with more emphasis on Galactic-Plane surveys. A full review covering all the pulsar surveys and the lessons learnt are beyond the scope of the thesis, however we give a brief summary in order to give context to the bulk of the scientific results of this thesis.

2.2.7 Previous Generation Surveys

The discovery of the first pulsar by Hewish et al. (1968) inspired the creation of multiple first-generation pulsar surveys by different radio telescopes around the world to discover as many similar exotic objects as possible. Some examples include the first Molonglo transit survey with the Molonglo Cross Telescope (MOL1; Large et al. 1968; Turtle & Vaughan 1968; Vaughan & Large 1970, 1972) which used a pen-chart recorder to store the signal of the receiver, the first Jodrell Bank survey (JB1) using the 76-m Lovell Telescope (Davies et al., 1970; Davies & Large, 1970; Davies et al., 1973) pushed this further by using a computer to automatically detect bright single-pulses. These surveys were typically conducted at a central frequency of 408 MHz with a bandwidth of 4 MHz using only two frequency channels. Despite, the relative success of these surveys, they were severely limited by today's standards and were only able to detect the brightest of pulsars. The clustering of pulsars close to the Galactic plane were remarkably already noticed in Large et al. (1968); Wielebinski et al. (1969). Both these studies reported an anomalous distribution with more pulsars south of the galactic plane and a zone of avoidance was proposed to explain the absence of pulsar within the plane itself. Today, we can explain these non-detections mostly due to the low sensitivity and coarse frequency resolution of the survey which hindered removing the effect of dispersion that is more pronounced close to the plane. In order to find weaker



```
                                   Search Information
Candidate:  523.94ms_Cand          RAJ2000 = 17:22:28.2000         DECJ2000 = −38:09:57.5780
Telescope:  Parkes                        Best Fit Parameters
Epochtopo = 55739.61712962963      DOFeff = 60.42  χ2red = 4.049  P(Noise) < 6.06e−25  (10.2σ)
Epochbary = N/A                    Dispersion Measure (DM; pc/cm3) = 689.341
Tsample         =  0.000256        Ptopo (ms) =  523.94418(78)     Pbary (ms) = N/A
Data Folded     =  16896000        P'topo (s/s) = −5.5(1.4)x10−9   P'bary (s/s) = N/A
Data Avg        =  1555            P''topo (s/s2) = 0.0(2.1)x10−12 P''bary (s/s2) = N/A
Data StdDev     =  16.38                  Binary Parameters
Profile Bins    =  64              Porb (s) = 28112.285180         e = 0.000000
Profile Avg     =  4.104e+08       a1sin(i)/c (s) = 1.250540       ω (deg) = 0.000000
Profile StdDev  =  8416            Tperi = 55739.48182929423
```

Figure 2.5: Discovery plot of PSR J1722-3809 which was found in the re-analysis of HTRU-S Lowlat Survey. We have marked the different sub-components of the plot that is typically used to identify if a signal is from a genuine pulsar. Plot marked as ‘1’ is the pulse profile of the signal, ‘2’ and ‘3’ are the Phase-Time and Phase Frequency plot respectively. Plot ‘4’ is the DM curve, Plot ‘5’ shows χ^2 vs P and Pdot which can be thought of as a fit for the best period and period derivative. Plot ‘6’ shows the signal in P-Pdot plane. More details about each of these plots can be found in section 4.2.6. This figure was created using the PRESTO software package.

pulsars, the first Arecibo survey (AR1; Hulse & Taylor 1974; Hulse 1975; Hulse & Taylor 1975a) was carried out using the 305-m Arecibo telescope at a central frequency of 430 MHz but with 32 channels across an 8 MHz bandwidth. They used Fourier techniques like the FFT and time-domain techniques like the FFA (Burns & Clark, 1969; Taylor & Huguenin, 1969) which are still in use today and found 40 new pulsars including the first binary pulsar ever to be found-PSR B1913+16 (Hulse & Taylor, 1975b). Arguably, the most successful pulsar survey of this era was the second Molonglo survey (MOL2; Manchester et al. 1978). MOL2 reported the discovery of 155 new pulsars single-handedly doubling the number of known pulsars at the time. Following MOL2, several significant milestones were reached including the discovery of the first MSP PSR B1937+21 (Backer et al., 1982) spinning at a rate of 1.56 ms. This was likely missed by earlier surveys due to their coarse time sampling rate which was in the order of few tens of milliseconds. This discovery inspired the second and third Arecibo surveys (AR2, Segelstein 1986) and (AR3, Nice et al. 1995), the Green-Bank short period survey using the 92-m telescope at Green Bank, (GB3, Stokes et al. 1985), the Parkes Southern survey (PKS70; Manchester et al. 1996; Lyne et al. 1998) which searched for MSPs with a much faster sampling rate and increased channelisation of the frequency bandwidth. All the surveys mentioned so far were conducted at low radio frequencies of 400-430 MHz. Surveys conducted at these frequencies are severely limited by selection effects like interstellar scattering (ν^{-4}), dispersion pulse broadening (ν^{-3}) and the high galactic background temperature close to the Galactic-Plane. Therefore, around this time, a number of attempts were made to do Galactic-plane surveys at higher frequencies. Some examples include the second Jodrell Bank Survey (JB2) (Clifton & Lyne, 1986), the Parkes 20-cm survey (Johnston et al., 1992) which had a sampling rate of 0.3 ms and a central frequency of 1434 MHz with a bandwidth of 80 MHz. The trade-off of observing at higher frequency is the decreased brightness of the pulsar signal and smaller telescope beam size (Area $\propto \nu^{-2}$), therefore requiring more observations to cover the same part of the sky.

2.2.8 The Parkes Multibeam Pulsar Survey

The Parkes Multibeam Pulsar Survey (PMPS, Manchester et al. 2001) is arguably the most successful pulsar survey ever to be conducted and is an important milestone in the history of pulsar astronomy. PMPS has so far found greater than 1000 pulsars with new pulsars still being discovered using new search techniques. Observations for this survey began in 1997, covering a galactic longitude $260^\circ < l < 50^\circ$ and galactic latitude $|b| < 5^\circ$ with an integration time of 35 min and a sampling rate of 250 µs. The success of this survey can be largely attributed to the then newly installed 20cm 13-beam Multi-Beam Receiver (MB-20, Staveley-Smith et al. 1996) at the focus-cabin of the Parkes Radio telescope. We refer the readers to Lyne (2008) for an interesting narration of the history and context surrounding the building of this receiver. It was originally intended for extragalatic H1 surveys like the HI Parkes All-Sky Survey (HIPASS) and Zone of Avoidance (ZOA), however due to the large number of beams and wide frequency bandwidth of 288 MHz, it was later extensively used for pulsar observations. Some notable achievements using PMPS data include the discovery of the relativistic binary PSR J1141-6545 (Kaspi et al., 2000) which has recently been used to infer Lense-Thirring precession (Venkatraman Krishnan et al., 2020) discovery of the first Rotating Radio Transients (RRATS, McLaughlin et al. 2006; Keane et al. 2010), three double neutron star (DNS) systems, PSR J1811-1736 (Lyne et al., 2000), PSR J1756-2251 (Faulkner et al., 2005) and PSRs J1753-2240 (Keith et al., 2009), and hundreds of young pulsars (Manchester et al., 2001; Morris et al., 2002; Kramer et al., 2003; Hobbs et al., 2004; Lorimer et al., 2006) which shed more light into the Galactic population of pulsars (Lorimer et al., 2006). Additionally, new pulsars were found a decade later in the reprocessing of PMPS. Some examples include 16 new pulsars found using a coherent acceleration search technique (Eatough et al., 2013), 24 pulsars with the Einstein@Home volunteer distributed computed project using the template-bank algorithm (Knispel et al., 2013). The first application of neural networks to classify pulsar candidates were also applied to the data from the PMPS survey (Eatough et al., 2010a).

2.3 The HTRU-S LowLat pulsar survey

HTRU-S Lowlat forms the bulk of the data that was analysed during this thesis. Therefore, we will expand more about this pulsar survey in the subsequent section.

2.3.1 Introduction and Survey goals

HTRU-S Lowlat is one part of the High Time Resolution Universe[9] (HTRU) which is an all sky, high time and frequency resolution blind survey aiming to find new pulsars and radio transient objects. HTRU-South (HTRU-S) covers the southern part of the sky and the observations were taken using the 64-m Parkes Radio Telescope in Australia at L-band (Keith et al., 2010). A corresponding survey that covers the northern part of the sky is currently ongoing with observations taken at the 100-m Effelsberg telescope in Germany (Barr et al., 2013). The HTRU survey consists of three parts covering different galactic latitude regions called High Latitude (HiLat), Med Latitude (MedLat) and Low Latitude (Lowlat). Broadly speaking, the aim of the high-latitude survey was to cover large parts of the sky at 64 µs time resolution to discover transient events. The Midlat survey was aimed at discovering bright isolated millisecond pulsars (MSPs) which would be useful for pulsar timing array projects. The aim of Lowlat is two-fold. First is to discover low luminosity pulsars deep in the Galactic plane and the second is to find relativistic binary pulsars which are expected to be found close to the galactic plane. A detailed list of the observational parameters of each component of the survey is given in table 2.1 and a visual representation of the sky coverage and overlap region can be found in figure 2.6. Other than the sky coverage and integration time, all other observational parameters and specifications of the survey are identical. For completeness, I describe the observational set-up here briefly. HTRU-South uses the same 13-beam Parkes 21-

---

[9] `https://sites.google.com/site/htrusouthdeep/home`

cm multibeam receiver (MB20) as used in previous Parkes surveys like the PMPS. The main technological improvement of HTRU over PMPS is the higher frequency and time resolution offered by the Berkeley Parkes Swinburne Recorder (BPSR) backend system (Keith et al., 2010). This increases the sensitivity of of HTRU towards MSPs as it increased the frequency resolution by a factor of 8 compared to the PMPS which minimises the effects of dispersive smearing. Additionally, the time resolution of 64 µs is a factor of four greater than the sampling time of 250 µs of the PMPS which mitigates the effects of aliasing for the higher harmonics of MSPs. Additionally, Lowlat also doubled the integration time with $T_{obs} = 1.2$h which makes it one of the most sensitive and deepest Galactic-Plane survey conducted till date. The sensitivity of the Lowlat survey to MSPs and Normal Pulsars is given in table 2.2. The observations were recorded as a polyphase filterbank in search-mode using dual-polarisation across 1024 channels at a channel bandwidth 391 kHz with a total usable bandwidth of 400 MHz. However, in practice due to RFI from the Tharuya 3 satellite, the usable bandwidth was reduced to 340 MHz, with 870 channels remaining. We refer the readers to (Keith et al., 2010) for a more detailed description of the observation set-up of the HTRU-South survey. The HTRU survey project is part of an international collaboration, with the primary responsibility of data-processing and follow-up shared between the Max-Planck-Institute für Radioastronomie (MPIfR) in Germany, the Swinburne University of Technology and the Australia Telescope National Facility (ATNF) in Australia, the Istituto Nazionale di Astrofisica (INAF) in Italy and the University of Manchester in the UK.

Table 2.1: Survey specifications for HTRU North and South Pulsar surveys. Table adapted from (Ng, 2014; Cameron, 2018)

	HTRU-North			HTRU-South		
Telescope	Effelsberg 100-m			Parkes-64m		
Receiver	7-Beam			MB20		
Backend	PFFTS			BPSR		
Peak gain, G (KJy^{-1})	1.5			0.735		
Receiver temperature, T_{rec} (K)	21			23		
Digitisation bit-rate, n_{bits}	8			2		
Number of polarisation summed, n_p	2			2		
Central frequency, f_c (MHz)	1360			1352		
Bandwidth, B (MHz)	240			340		
Number of channels, n_{chan}	512			870[a]		
Channel bandwidth, Δf_{chan} (kHz)	468.75			390.625		
Sampling time, t_{samp} (µs)	54.6			64		
Region of declination, $\delta(^\circ)$	$\delta >$ -20			$\delta <$ +10		
Survey region	HiLat	MedLat	Lowlat	HiLat	MedLat	Lowlat
Sky region in Galactic longitude, $l(^\circ)$	-	-	-	-	$-120 < l < 30$	$-80 < l < 30$
Sky region in Galactic latitude, $b(^\circ)$	$\|b\| > 15$	$\|b\| < 15$	$\|b\| < 3.5$	-	$\|b\| < 15$	$\|b\| < 3.5$
Integration time, t_{obs}	90	180	1500	270	540	4300
Characteristic minimum sensitivity,[b] S_{min} (mJy)	0.61	0.34	0.13	0.4	0.3	0.13
Mean sky temperature, T_{sky}[c] (K)	5	8	11	1.0	2.5	7.6
Number of beams, N_{beams}	1066135	375067	87395	443287	95056	15990
Number of samples per beam, n_{samp}	1.6×10^6	3.3×10^6	27.4×10^6	$\sim 2^{22}$	$\sim 2^{23}$	$\sim 2^{26}$
Data size per beam (GB)	0.8	1.6	13.4	1.0	2.0	16.6
Total data size (TB)	818	576	1118	435	190	263

[a] Originally contains 1024 channels but due to RFI filters only 870 channels are usable.

[b] Calculated for a pulsar with spin period of a few ms and a duty cycle of 30 per cent.

[c] Values of Tsky extrapolated from the 408-MHz sky-map presented in Haslam et al. (1981).

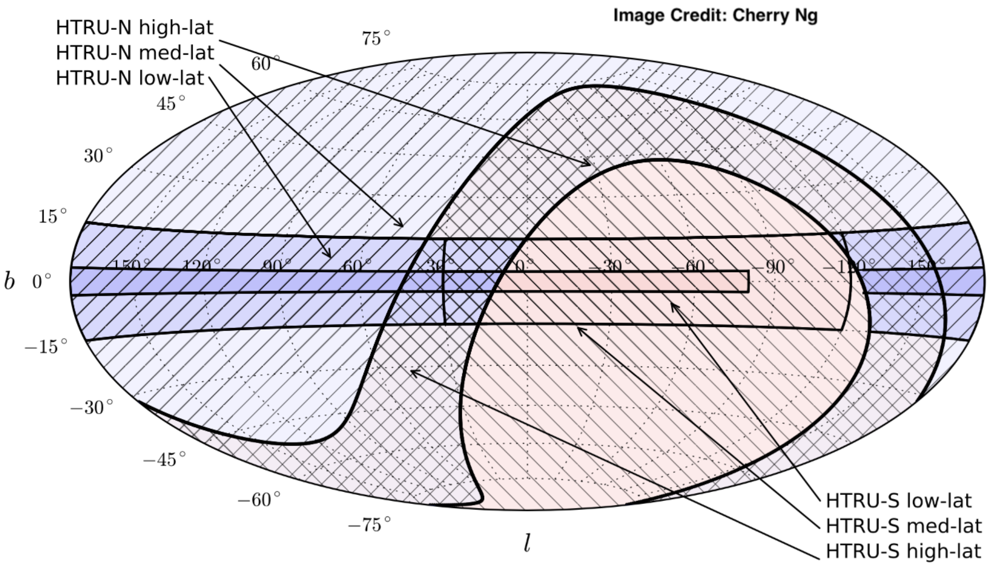


Figure 2.6: Sky coverage of HTRU pulsar survey. Red regions marked with '\' shows the part of the sky covered by the southern survey whereas blue regions marked with '/' indicate the region covered by the northern survey. A full list of the survey specifications can be found in table 2.1. Figure taken from Ng (2014)

Table 2.2: Characteristic minimum detectable flux density $S_{1400,min}$ also known as sensitivity of the HTRU-S Lowlat pulsar survey. The minimum, mean and maximum values of sensitivity are reported for populations of MSPs (P $<$ 30 ms) and Normal pulsars with different pulsar duty cycles. Table adapted from (Ng, 2014)

	MSPs		Normal Pulsars	
	δ(%)	$S_{1400,min}$(mJy)	δ(%)	$S_{1400,min}$(mJy)
min	0.28	0.013	0.014	0.0030
mean	11.54	0.092	4.21	0.053
max	65.31	0.35	57.29	0.29

2.3.2 Discovery Highlights

The HTRU-S as a whole has already discovered at least 230 new pulsars (Keith et al., 2010; Bailes et al., 2011; Bates et al., 2011; Keith et al., 2012; Burgay et al., 2013; Ng et al., 2014; Bates et al., 2015). Of this, the first pass processing of the Lowlat portion of the survey contributed 100 new discoveries. A full list of these pulsars and initial timing solutions can be found in Ng et al. (2015); Cameron et al. (2020). The former study also includes a detailed description of the observing and search strategy. The flagship discovery of Lowlat so far has been the detection of one of the most relativistic binary pulsars PSR J1757-1854, a DNS in a highly eccentric (e∼0.6) 4.4-hour orbit (Cameron et al., 2018), which was used to perform three independent tests of GR, and future timing observations are expected to measure Lense-Thirring precession. Some additional highlights include PSR J1653-4518, a 1.5-year eclipsing binary, PSR J1812-15, a 1-s pulsar in a 20.3-hr circular orbit around a light companion which also exhibits variable emission switching between on and off states (Cameron et al., 2020), PSR J1755-2550, a young pulsar with a massive, compact companion (Ng et al., 2018), PSR J1759-24, an eclipsing binary system with a long orbital period (Ng et al., 2015). Some other notable achievements of the general HTRU survey include the discovery of a radio loud magnetar, PSR J1622-4950 (Levin et al., 2010), PSR J1719-1438, a 'Pulsar-Planet' binary (Bailes et al., 2011). Many of these pulsars are still being followed up by various telescopes around the world and are expected to produce more exciting scientific results in the future. Despite the successes of HTRU so far, as shown from the example of PMPS many pulsars can still be discovered by employing novel search algorithms. The first pass processing of Lowlat used a segmented-acceleration search pipeline (Ng et al., 2015) which is sensitive to pulsars with a constant acceleration during an observation. Therefore, relativistic binary pulsars with a changing acceleration (also called jerk) could have been missed in past analysis. In the next chapter we will report results from the re-processing of Lowlat with novel algorithms like the template-bank algorithm which searches directly over Keplerian Parameters. This improves our sensitivity towards relativistic binary pulsars at a higher computational cost. We describe the algorithm, the CPU and GPU pipeline developed as a part of this thesis and the new discovery of 20 new pulsars including a MSP which has been confirmed to be in a binary orbit. We also additional discuss a novel semi-supervised machine learning solution which has been incorporated to the Lowlat pipeline to reduce the number of pulsar candidates that need to be inspected by eye.

Chapter 3

Keplerian-Parameter Searches for PSR-BH Binaries in Circular Orbits in the HTRU-South Lowlat Survey

---

Contents

3.1 Motivation: Why search for PSR-BH Binaries?

To test gravity in extreme conditions requires either a high gravitational potential or for the compact objects involved to have relativistic velocities. Pulsars can be thought of as high precision clocks attached to point masses, therefore the discovery and subsequent timing of a pulsar orbiting a black hole will provide a unique opportunity for testing GR and alternate theories of gravity in the strong-field limit near the vicinity of a black-hole. The discovery of such systems opens up a parameter space that cannot be explored by Solar System tests, Double-Neutron Star Binaries or PSR-WD binaries. This is shown in the left hand side of figure 3.1 where we show the Gravitational self-energy ε expressed in terms of its rest mass energy as $\varepsilon \sim -GM/Rc^2$, where G is the Gravitational constant, M is the mass of the compact object, R is its radius and c is the speed of light in vacuum is plotted against its orbital size. The lower right half of this diagram is excluded as it implies a compact object with orbit smaller than its size. We see that $\varepsilon \sim -10^{-10}$ for Earth, $\varepsilon \sim -10^{-6}$ for the Sun, $\varepsilon \sim -0.2, 0.5$ for a NS and BH respectively. Alternative theories of gravity like scalar-tensor theories (Damour & Esposito-Farese, 1993; Damour & Esposito-Farèse, 1996) predict the properties of gravity to be dependent on the Gravitational self-energy ε. Therefore, by timing objects having different gravitational self-energies in an orbit with a common external gravitational potential can help us test the Strong Equivalence Principle (SEP) which is a prediction of GR but is expected to be violated in alternate theories of gravity (Deruelle, 2011). The most stringent limits for violation of SEP are currently provided by timing a PSR-WD system (Archibald et al., 2018; Voisin et al., 2020). However, as shown in the right hand side of figure 3.1 where Schwarzschild radius of the compact object is plotted against its gravitational mass, PSR-BH binaries are expected to have larger strong-field effects and therefore by testing SEP in these regimes could either lead to a violation of GR or put even more stringent limits on alternate theories of gravity. Additionally, such discoveries will almost certainly improve our understanding of stellar evolution of massive progenitor stars. PSR-BH binaries are also expected to improve our understanding of BH-Physics (Paczynski & Trimble, 1979). As discussed earlier in section 1.8, by continuously monitoring the radio pulsar in timing mode we can measure the orbit's Post-Keplerian Parameters. Measurement of two Post-Keplerian Parameters would give us as a unique measurement for the mass of the pulsar and the BH assuming GR. The spin and quadrupole moment of the BH can also be determined through pulsar timing by measuring the effects of relativistic spin-orbit coupling and classical spin-orbit coupling respectively. These measurements could be used to directly test the “No-hair" theorem and the Cosmic-Censorship Conjecture. (Wex & Kopeikin, 1999; Kramer et al., 2004).

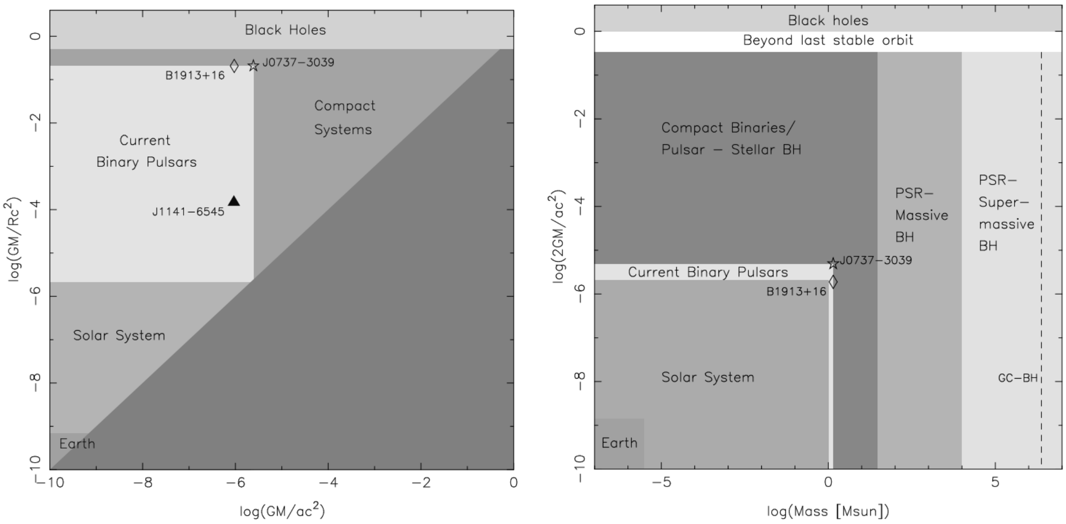


Figure 3.1: Gravitational Regime probed by PSR-BH Binaries. Figure from Kramer et al. (2004). See main text for discussion.

3.1.1 Formation

There are two main formation channels hypothesised to create a PSR-BH binary. One is from the evolution of massive binary stars in the Galactic field analogous to the formation of DNS systems discussed earlier in section 1.6.3. The alternative is the dynamical formation scenario. This occurs in areas of high stellar density like Globular Clusters (GCs) or the Galactic Centre which increases the probability of three-body interactions. These interactions can cause a NS to gain a companion by exchanging with a primordial binary and subsequently be spun up to become an MSP (Hills, 1976; Sigurdsson & Phinney, 1995; Ivanova et al., 2008; Clausen et al., 2014). In this section, we will limit our discussion to the former scenario which is relevant for the processing of Galactic-Plane observations undertaken during this thesis. In the standard binary stellar evolution model (Yungelson & Portegies Zwart, 1998; Voss & Tauris, 2003), the more massive primary star will evolve first forming a black hole followed by the secondary star which forms a radio pulsar at a later epoch. Since the PSR is formed second, no spin-up is possible hence we expect to find a slow spinning normal pulsar orbiting a black hole in a wide eccentric orbit in this formation channel. Alternatively, a BH+MSP binary can be formed by the so-called reversal mechanism where under certain conditions the pulsar is formed first and is later spun-up by accretion during the red giant phase of the secondary star (Sipior et al., 2004; Pfahl et al., 2005). Some examples of binary pulsars which had a similar evolutionary history are PSR B2303+46 and PSR J1141-6545, both of which are pulsar-white dwarf binaries in which the neutron star was born after the white dwarf (Tauris & Sennels, 2000; Davies et al., 2002). Discovering BH+MSP binary systems are more desirable as MSPs tend to be more precise timers than normal pulsars (e.g. Verbiest et al. 2009). Population synthesis work by Chattopadhyay et al. (2021) has shown that binaries where the NS is formed first are 3-4 orders of magnitude rarer compared to the scenario in which the BH is formed first. A caveat to add here is that all these predictions depend on the formation rate of PSR-BH binaries in the Galaxy which is not well understood. Narayan et al. (1991) estimated that the number and formation rate of BH-NS binaries in the Galaxy are comparable to the corresponding estimates of DNS systems. However, Pfahl et al. (2005) estimated that fewer than 10 of PSR-BH binaries currently reside in the Galactic disk less than 0.1%-1% of the number of DNS systems. Kruckow et al. (2018) using their population synthesis simulations estimated that mixed systems like NS-BH binaries are the most commonly formed double compact objects. Therefore, this remains an open question and will likely only get resolved by more sensitive observations and/or detections of such systems. However, despite several observational efforts spanning across decades covering a wide range of frequencies, none have been found so far. Using the prior information available from previous searches and based on the results of population synthesis studies mentioned above, we decided to undertake two separate search strategies. One targeting MSP+BH binaries and another targeting BH+normal PSR binaries in more compact orbits. We will expand more on this in section 3.2.4.

3.1.2 Limitations of current search strategies

Finding new binary pulsars in compact orbits is challenging because the apparent period of the pulsar changes during the course of observation due to Doppler modulation as the pulsar moves around its orbit. If this is unaccounted for, then the signal will get smeared out in the FFT. If we know the exact line of sight radial velocity of the pulsar, then we can add or subtract appropriate delays in our timeseries in order to recover the signal from the observations. However, since before discovering a pulsar we do not know the velocity, we need to search through a range of possible values. With no prior information, this would lead to a five parameter search for each of the Keplerian Parameters in addition to the usual DM trials and period trials in the FFT for a standard pipeline. This has historically been computationally unfeasible. The most commonly used technique is to first order assume that the line of sight acceleration is constant during the course of an observation i.e $V = a \times t$, which adds just one extra parameter a that needs to be searched. Correcting for acceleration can be done both in the time-domain (Johnston & Kulkarni, 1991) and frequency domain (Ransom et al., 2002). While this technique has been used successfully to discover several binary pulsars the assumption of constant acceleration breaks down when we sample a large fraction of the orbit in an observation i.e $P_{\mathrm{orb}} \leq 10\, T_{\mathrm{obs}}$[1] (Ransom et al., 2003; Ng et al., 2015). Therefore, shorter observations are more desirable for finding binary pulsars. However, the sensitivity of a pulsar signal scales as $\mathrm{T_{obs}}^{1/2}$, therefore pipelines built on the acceleration-search algorithm are limited to finding binaries in long orbits. In practise a mixed strategy is usually adopted with mild acceleration trials done on the full-length observation and higher acceleration trials done on “segments" of the data. This was also the processing strategy used in the first-pass processing pipeline of HTRU-S Lowlat called a partially coherent segmented-acceleration search pipeline (Ng et al., 2015). In

---

[1] This will often be referred to in this thesis as the 10% rule.

order to address these drawbacks a new GPU pipeline was built as a part of this PhD thesis built on previous works done by Knispel (2011); Knispel et al. (2013); Allen et al. (2013) and the gravitational-wave search community. This approach uses the template-bank algorithm to build orbital templates for potential binary pulsar signals by directly searching over Keplerian Parameters assuming a circular orbit. Over the next few sections, we will describe this technique and demonstrate the extra sensitivity gained by our pipeline by comparing it to standard pulsar search techniques used in literature for simulated PSR-BH observations. We will also describe our results of applying this technique on data from HTRU-S Lowlat which is the most sensitive Galactic-Plane survey at L-Band for PSR-BH systems in the southern hemisphere.

3.2 Methods: Coherent Search over Keplerian Parameters for Circular Orbits using the template bank algorithm

The template bank construction method described here is similar to the technique described in Chapter 5 which is applicable for the general case of elliptical orbits. Here, we expand on the special case of circular orbits, and its application on observations of HTRU-S Lowlat data.

3.2.1 Signal Model for Circular Orbits

We first start by defining the rotational phase of a binary pulsar which acts as our signal model. We make the following assumptions in our signal model. The orbit of the pulsar is Keplerian.[2] We know the exact location of the source on the sky, therefore Doppler phase drifts caused due to the detector motion can be removed and the signal is monochromatic at the spin-frequency of the pulsar f_0 i.e we ignore spin-down effects of the pulsar. Therefore, the rotational phase Φ for the fundamental mode of the signal emitted by a uniformly rotating pulsar in a circular orbit observed from solar-system barycenter is

$$\Phi(t) = 2\pi f_0 t + \phi_D(t), \tag{3.1}$$

where t is the barycentric time, f_0 is the spin-frequency of the pulsar and $\phi_D(t)$ is the Doppler phase correction factor due to the pulsar being in a circular orbit, given by (see section 2B of Dhurandhar & Vecchio 2001 for the full derivation of this expression and section 5.2.1 for the general case of elliptical orbits.),

$$\phi_D(t;\Lambda) = \frac{2\pi f_0 a \sin\varepsilon}{c} \sin(\Omega t + \alpha). \tag{3.2}$$

where, ε is the inclination angle of the pulsar orbit, Ω is the angular velocity of the pulsar which is related to its orbital period as $\Omega = \frac{2\pi}{P_{orb}}$, α is the initial orbital phase

---

[2] This assumption is based on the typical timescales of search-mode observation of radio pulsars which is in the order of few hours. For long observations lasting months or years, additional parameters including Post-Keplerian and spin-down parameters need to be added to the phase model.

and c is the speed of light in vacuum. Combining equations 3.1 and 3.2, we get

$$\phi(t;\Lambda) = 2\pi f_0\big[t + \tau\sin(\Omega t + \alpha)\big]. \tag{3.3}$$

where $\tau = \frac{a\sin(\varepsilon)}{c}$ is the projected semi major axis in light seconds, $\Lambda = \{f_0, \Omega, \tau, \alpha\}$ is a tuple of the Keplerian orbital parameters and spin frequency of the pulsar signal. Since, *a priori* these values are unknown, we need to search through a wide range of trials, each trial can be thought of as a "template" of the signal that is under consideration. Therefore, in order to do a blind coherent search for circular orbits we need to search through a four dimensional parameter space of $f_0, \Omega, \tau, \alpha$. We then do a matched filtering process by convolving the detector output with various templates of the pulsar signal. This is a widely used technique in Gravitational-wave data analysis (for e.g. Owen & Sathyaprakash 1999; Abbot et al. 2007, 2009). In order to sample efficiently in this four-dimensional grid, we need a measure of distance in the parameter space. The distance can be approximately calculated by computing the parameter space metric (Owen, 1996) which will be defined in the next section. We will then discuss some of the limitations of this metric approximation and describe a technique that overcomes this.

3.2.2 Definition of Parameter Space Metric and Mismatch

Before we define the metric we first start by defining a coherent detection statistic. We know that the total time-domain radio intensity is a sum of instrumental, environmental noise $N(t)$ and harmonic sums of a pulsar signal. This can be written as:

$$s(t;\Lambda) \equiv N(t) + \sum_{n=1}^{\infty} s_n(t;\Lambda), \tag{3.4}$$

where the intensity of each harmonic is given by

$$s_n(t;\Lambda_1) \equiv \Re\big[\mathcal{A}_n \mathrm{exp}[in\Phi(t;\Lambda]\big], \tag{3.5}$$

where $\mathcal{A}_n$ are the complex amplitudes of the harmonics of the signal. We then define a coherent detection statistic $P_n(\Lambda, \Lambda')$ for the nth harmonic computed using the radio intensity correlated with the nth normalised signal template currently being searched $\exp[-in\Phi(t;\Lambda)]$. This is the detection statistic recovered from a pulsar signal with true parameters at Λ with a template at Λ'.

$$P_n(\Lambda,\Lambda') = \left|\frac{1}{T}\int_0^T dt\, s(t;\Lambda)\exp\left[-in\Phi(t;\Lambda')\right]\right|^2. \tag{3.6}$$

Thresholding the detection statistic helps us minimize the false-alarm probability at a fixed rate of false-alarm probability Allen et al. (2002). In the absence of a pulsar signal, the $N(t)$ term dominates and the detection statistic term is proportional to the instrumental noise whereas in the presence of a strong signal the $N(t)$ term can be neglected and the expectation value of the detection statistic then becomes,

$$\langle P_n(\Lambda, \Lambda')\rangle \approx \left|\frac{A_n}{2}\right|^2 \left|\frac{1}{T}\int_0^T dt\, \exp[in(\Phi(t;\Lambda) - \Phi(t;\Lambda'))]\right|^2 . \tag{3.7}$$

A priori we do not know the pulse-profile of the signal, therefore we construct detection statistics for up-to $32(2^5)$ harmonic sums i.e $P_0, P_1 ... P_5$ and we equally weight them assuming a Dirac-delta pulse profile. This can be written as:

$$S_L = \sum_{n=1}^{2^L} P_n. \tag{3.8}$$

where 2^L is the number of harmonic sums performed.

Using equation 3.7, we can define the mismatch (fractional loss of detection statistic in our case signal to noise ratio) between two points as

$$m(\Lambda, \Lambda') = 1 - \frac{P_n(\Lambda, \Lambda')}{P_n(\Lambda, \Lambda)} \approx g_{\alpha\beta}\, \Delta\Lambda^\alpha \Delta\Lambda^\beta + \mathcal{O}(\Delta\Lambda^3), \tag{3.9}$$

where a mismatch of one implies complete loss of signal and a mismatch of zero implies a perfect recovery of the signal, $g_{\alpha\beta}$ is the metric tensor, α and β correspond to points in the four-parameter space described earlier and we adopt the Einstein summation convention where repeated indices are summed over. For small deviations of the parameter space coordinates, $g_{\alpha\beta}$ can be calculated by Taylor expansion as (section 8.5 of Knispel (2011) for the full derivation)

$$g_{\alpha\beta} = \langle \partial_\alpha \Phi \partial_\beta \Phi \rangle_T - \langle \partial_\alpha \Phi \rangle_T \langle \partial_\beta \Phi \rangle_T, \tag{3.10}$$

where the angle brackets represent time average of a function $G(t)$

$$\langle G(t) \rangle_T \equiv \frac{1}{T}\int_0^T dt G(t). \tag{3.11}$$

3.2.3 Frequency-Projected Metric

In principle, we can construct a template-bank for the four dimensional parameter space of $\Lambda = \{f_0, \Omega, \tau, \alpha\}$, however since we have an efficient algorithm (FFT; Cooley & Tukey 1965) to search in f_0, in practise for computational reasons, we calculate a frequency projected metric tensor which is defined for the maximum spin-frequency of the of the harmonic of the signal f_{max} that our search is sensitive towards. This can be thought of as a slice across our parameter space for a given value of f_0. Our search analysis then becomes a Cartesian product of a three-dimensional orbital template-bank $\Lambda_{\text{orb}} = \{\Omega, \tau, \alpha\}$ along with a uniformly spaced grid in the frequency axis. This can be calculated as:

$$\gamma_{\alpha\beta} = g_{\alpha\beta} - \frac{g_{f\alpha} g_{f\beta}}{g_{ff}}, \tag{3.12}$$

where the repeated indices are summed over. Therefore, for each orbital template, we resample our timeseries to the binary system's barycenter by applying a time-domain resampling algorithm (described in section 5.5.1) and then calculate an FFT. For circular orbit searches, the number of orbital templates scales to the third power of the spin-frequency $\left(\mathrm{N}_{\text{templates}} \propto \mathrm{f}_0^3\right)$. For pulsars with spin-frequency $f_0 > f_{\max}$, we have a lower sensitivity i.e higher average mismatch and conversely for spin frequency $f_0 < f_{\max}$, we have increased sensitivity i.e on average a lower mismatch value.

3.2.4 Constraining the Parameter Space for PSR-BH Binary Searches

Before, we construct a template-bank, we need to decide the range of the parameter space ($\Lambda = \{f_0, \Omega, \tau, \alpha\}$) for our search analysis. Our goal while designing the search range, was to explore an astrophysically interesting parameter space that is yet to be investigated. One of the primary goals of the HTRU-S Lowlat survey was to discover new relativistic binary pulsars. The anticipated discovery of the first pulsar-black hole binary is considered a holy grail in astrophysics. Such systems are expected to contain an unrecycled pulsar because the black hole forms first, and are expected to have highly eccentric orbits. A full five keperlian parameter search for PSR-BH binaries in eccentric orbits for a survey like LOWLAT is still not computationally feasible. We discuss this in more detail later in section 5.2.1. Since a pulsar in a highly eccentric orbit spends most of its time away from periastron, a circular orbit search can still be sensitive to parts of its orbital phase and put limits for the existence of such systems. Therefore we constrain our search-range using our astrophysical prior information for likely orbital and spin-frequency parameters of PSR-stellar Mass BHs and the available computational resources that can be used for our analysis. We exclude parameter space regions that have already been investigated by previous search analysis. The first-pass processing of HTRU-S Lowlat used a segmented-acceleration search pipeline (Ng et al., 2015; Cameron et al., 2020) with the maximum line of sight acceleration for each segment chosen in order to be sensitive towards a PSR-BH system with a minimum orbital period $P_{\text{orb}} = 10\,T_{\text{obs}}$, BH mass of 10 $\mathrm{M}_\odot$ in a circular orbit ($e = 0$) with an edge-on inclination ($\varepsilon = 90°$). The acceleration ranges used for each segment of HTRU-S Lowlat can be found in table 3.1. The minimum orbital period threshold comes from the previously defined 10 % rule due to the assumption of constant "acceleration". The advantage of using the template-bank algorithm is that we do not have this lower orbital period constrain. In principle, it can be used to search for circular orbits of arbitrary orbital periods assuming computational feasibility. Given these prior searches, we place an upper limit on the orbital period in our template with $P_{\text{b,max}} = 12\text{h}$ for a 72-minute observation assuming that binaries with longer orbital periods would have already been discovered by the acceleration-search pipeline. We focused our reprocessing efforts on the 72-minute full-length observations given the extra computational trials required to probe our four dimensional parameter space of $\Lambda = \{f_0, \Omega, \tau, \alpha\}$ and we leave the processing of segmented observations using the template-bank algorithm for future reprocessing efforts. Our lower limit for orbital period was decided solely based on the

Table 3.1: A summary of the acceleration search ranges probed by the first-pass processing pipeline of HTRU-S Lowlat using the Segmented Acceleration-Search Pipeline. It includes the minimum orbital period of a binary pulsar the search is sensitive towards and the total number of segments (s) in each group. Table adapted from Cameron et al. (2020)

Segment	s	T_{obs} (s)	$P_{\mathrm{b,min}}$ (h)	$\lvert a_{min}\rvert$ ($\mathrm{m}s^{-2}$)	$\lvert a_{max}\rvert$ ($\mathrm{m}s^{-2}$)
Full-length	1	4320	12	0	1
Half-length	2	2160	6	0	200
Quarter-length	4	1080	3	200[a]	500
Eighth-length	8	540	1.5	0	1200

[a] A minority of the HTRU pointings were processed with $|a_{min}| \leq 200ms^{-2}$ in the s=4 segment.

amount of computation resources available.

In section 3.1.1, we discussed that assuming the standard binary stellar evolution model, Normal Pulsar+BH binaries are 3-4 orders of magnitude more likely than MSP+BH binaries (Chattopadhyay et al., 2021). Additionally, the number of required orbital templates scales as $\left(\mathrm{N}_{\mathrm{templates}} \propto \mathrm{f}_0^3\right)$. Therefore, we designed two separate searches in order to maximise our chances for detecting these separate populations of PSR-BH binaries. We call this a **"Recycled PSR-BH Binary Search"** (RPB) and a **"Non-Recycled PSR-BH" Binary Search** (NRPB). For our RPB search, we chose an f_{max} = 77Hz i.e $\mathrm{P}_{\mathrm{spin,min}} \sim 13.0$ms and for the NRPB search, we chose an f_{max} = 10Hz i.e $\mathrm{P}_{\mathrm{spin,min}} = 100.0$ms. Given the lower computational load for our NRPB search, due to the low value of f_{max}, we designed our template-bank to cover the orbital period range between $P_{\mathrm{orb}} = 3-12$h whereas for the RPB searches we cover $P_{\mathrm{orb}} = 6-12$h. The computational time for an FFT scales as $\mathcal{O}(n \log n)$, where n is the number of samples in the timeseries. Each LOWLAT observation spans 72-minutes and have a native time resolution of 64 µs. Therefore, once we dedisperse the data, this creates a timeseries of 2^{26} samples. Calculating an FFT of 2^{26} samples for each orbital template and DM trial for every observation in LOWLAT is computationally very expensive. Additionally, any detections in the FFT would also have to be folded on observations at the native time resolution which adds up to the total processing time of the survey. Therefore, for purely practical reasons for the RPB search, we downsampled the data by a factor of 4 to 256 µs which gives us a Nyquist frequency of 1.95 kHz and access to up-to 16 harmonics of a hypothetical 77 Hz signal, and for the NRP search we downsampled the data by a factor of 16 to 1024 µs, Nyquist frequency = 488.28 Hz. This gives us access to up-to 32 harmonics of a hypothetical 10 Hz signal below the Nyquist limit. The orbital and spin-parameters covered for each search is given in table 3.2 and the DM range and frequency range probed by our pipeline can be found in table 3.3.

Table 3.2: A summary of the orbital and spin period range used to create our template bank. We placed no constrains on the initial orbital phase and our maximum projected semi-major values were chosen based on kepler's laws for an orbit with inclination angle of 90 degrees and masses of pulsar and companion given below.

Search Type	t_{int} (h)	Min. P_{b} (h)	Max. P_{b} (h)	$\mathrm{Mass}_{\mathrm{pulsar}}$ ($M_{\odot}$)	$\mathrm{Mass}_{\mathrm{companion}}$ ($M_{\odot}$)	$\mathrm{P}_{\mathrm{spin,min}}$ (ms)
Recycled PSR-BH Binary Search (RPB)	1.2	6.0	12.0	1.4	8.0	13.0
Non-Recycled PSR-BH Binary Search (NRPB)	1.2	3.0	12.0	1.4	15.0	100.0

Table 3.3: Search Configuration for both our Lowlat searches displaying the Maximum DM searched, the frequency range probed in our FFT, and the total beams processed. DM tolerance DM_{tol} and DM pulse width W_{int} control the spacing of the DM trials.

Search Parameter	Lowlat
$\mathrm{DM}_{\mathrm{max}}$ ($\mathrm{pc\,cm^{-3}}$)	3000
$\mathrm{N}_{\mathrm{DMtrials}}$	1876 (RPB), 1126 (NRPB)
$\mathrm{DM}_{\mathrm{tol}}$	1.11
$\mathrm{W}_{\mathrm{int}}$(µs)	64
$\mathrm{N}_{\mathrm{harmonics}}$	16
$\mathrm{f}_{\mathrm{min}}$ (Hz)	0.1
$\mathrm{f}_{\mathrm{max}}$ (Hz)	1100.0
$\mathrm{N}_{\mathrm{beams}}$	14,488
Fraction processed	87.1 %

3.2.4.1 Constraint on projected semi-major axis

Our upper limit on the projected semi-major axis depends on the orbital period, pulsar and companion masses we want our search to be sensitive towards. Using Kepler's 3rd law, we know that

$$a^3\Omega_{\mathrm{orb}}^2 = \frac{Gm_{\mathrm{comp}}^3}{(m_{\mathrm{pulsar}} + m_{\mathrm{comp}})^2}, \tag{3.13}$$

where a is the orbital radius, G is the gravitational constant, c is the speed of light in vacuum, m_{pulsar} and m_{comp} is the mass of the pulsar and companion respectively. Rearranging equation 3.13 and writing in terms of projected semi-major axis in light-seconds $\tau = \frac{a \sin\varepsilon}{c}$, we get

$$\tau = \sin\varepsilon \times F(m_{\mathrm{comp}}, m_{\mathrm{pulsar}}) \times \Omega_{\mathrm{orb}}^{-\frac{2}{3}}. \tag{3.14}$$

where $F(m_{\mathrm{comp}}, m_{\mathrm{pulsar}})$ is a mass-dependent scaling factor given as:

$$F(m_{\mathrm{comp}}, m_{\mathrm{pulsar}}) = \frac{G^{\frac{1}{3}} m_{\mathrm{comp}}}{c\,(m_{\mathrm{pulsar}} + m_{\mathrm{comp}})^{\frac{2}{3}}} \tag{3.15}$$

For a given orbital period and inclination angle, the projected radius τ is maximum when the pulsar mass m_{pulsar} is minimal and the companion mass m_{comp} is maximal. τ is also directly proportional to the sine of the inclination angle ($\theta = \sin\varepsilon$) with the maximal value obtained for the edge-on case i.e $\varepsilon = 90°$. When $\varepsilon = 0°$ i.e the orbit is face-on with respect to the observer, $\tau = 0$ and no Doppler Modulation is observed. Using these properties, we calculate the constraint on τ as:

$$0 \leq \tau \leq \theta \times F(m_{\mathrm{comp,max}}, m_{\mathrm{pulsar,min}}) \times \Omega_{\mathrm{orb}}^{-\frac{2}{3}}, \tag{3.16}$$

where $m_{\mathrm{comp,max}}$ is the maximum companion mass, $m_{\mathrm{pulsar,min}}$ is the minimum pulsar mass. For both our searches we set the minimum pulsar mass $= 1.4M_{\odot}$. The θ parameter can be used to constrain the maximum inclination angle we want our search to be sensitive to, however we chose $\theta = 1$, so that we are fully sensitive even in the worst-case edge-on scenario with maximal Doppler modulation. For our RPB search we set the maximum companion mass $m_{\mathrm{comp,max}} = 8M_{\odot}$ and for our NRPB search we set $m_{\mathrm{comp,max}} = 15M_{\odot}$. For the initial orbital phase α we placed no constraint $\alpha = [0, 2\pi)$ radians.

Orbital inclinations are expected to be distributed randomly in nature. Thus the orbital inclination unit vectors can be thought of as distributed randomly on the surface of a unit sphere, thus making it uniform in $\cos\varepsilon$. We can calculate the fraction p of binary systems with inclination angle between 0 and $\varepsilon_{\mathrm{max}}$ as:

$$p = \int_0^{\varepsilon_{\mathrm{max}}} d\varepsilon \sin\varepsilon = 1 - \cos\varepsilon_{\mathrm{max}} = 1 - \sqrt{1 - \sin^2\varepsilon_{\mathrm{max}}} = 1 - \sqrt{1 - \theta^2} \tag{3.17}$$

As shown in equation 3.14, the projected semi-major axis τ depends on the mass ratio and the inclination angle for a given orbital period. Therefore, since we selected $\theta = 1$, if the mass of the companion $m_{\mathrm{comp}} \leq m_{\mathrm{comp,max}}$ and the mass of the pulsar $m_{\mathrm{pulsar}} \geq m_{\mathrm{pulsar,min}}$, then we can detect the system for any orbital phase and inclination angles assuming the flux density of the binary system is within our sensitivity limits. Additionally, we also retain a non-zero sensitivity for higher companion masses and lower pulsar masses. For any arbitrary pulsar and companion mass, using equation 3.13, the orbital radius $a = F(m_{\mathrm{comp}}, m_{\mathrm{pulsar}})\Omega_{\mathrm{orb}}^{-2/3}$. Substituting this expression to the left hand side of equation 3.16, we get :

$$\theta(m_{\mathrm{pulsar}}, m_{\mathrm{comp}}) = \theta_0 \frac{F(m_{\mathrm{comp,max}}, m_{\mathrm{pulsar,min}})}{F(m_{\mathrm{comp}}, m_{\mathrm{pulsar}})}, \tag{3.18}$$

where $\theta_0 = 1$ is our chosen bound on the sine of the inclination angle for our search. Therefore, using equation 3.17, we can calculate the fraction p as a function of pulsar and companion mass which is detectable due to favorable orbital inclination as:

$$p(m_{\mathrm{pulsar}}, m_{\mathrm{comp}}) = 1 - \sqrt{1 - \theta^2(m_{\mathrm{pulsar}}, m_{\mathrm{comp}})}. \tag{3.19}$$

For RPB search $p(m_{\mathrm{pulsar}}, m_{\mathrm{comp}})$ is shown in figure 3.2. Region marked in "white" is detectable for any inclination angle and orbital phase whereas for the rest of the diagram we retain only partial detectability depending on the orbital inclination angle. For the NRPB search, $p(m_{\mathrm{pulsar}}, m_{\mathrm{comp}})$ is shown in figure 3.3. Due to the higher companion mass limit of this search, we can detect a wider range in this plane for any inclination angles and orbital phases. However, the caveat is that this is true only for unrecycled pulsars with spin-period $P_{\mathrm{spin}} \geq 100$ms and for binaries in circular orbits.

3.2.5 Orbital Template Bank Construction

In this section we describe our technique for constructing an orbital template bank which was used for the reprocessing of HTRU-S Lowlat observations. As mentioned earlier, we construct orbital templates for the binary parameter space $\Lambda_{\mathrm{orb}} = \{\Omega, \tau, \alpha\}$ along with a uniform grid in the spin frequency axis f. Since *a priori*, we do not know the orbital and spin parameters of the pulsar, we search through various templates of a hypothetical pulsar signal. A combination of many such templates form a template-bank. Our template-bank construction method is similar to previous works by Knispel (2011); Allen et al. (2013) and Knispel et al. (2013). These works were in turn based on the extensive research done by the gravitational wave search community. (for e.g. Owen 1996; Owen & Sathyaprakash 1999; Abbot et al. 2007, 2009; Messenger et al. 2009; Harry et al. 2009). Prix (2007) showed that generating a template bank is an instance of the sphere covering problem and techniques from this field can be used to construct optimal template banks. There are two characteristic quantities for a template-bank-coverage η and mismatch m. The classical approach of creating a lattice-based template-bank has been to cover every point in the parameter space with a template i.e $\eta = 1$. However, this quickly becomes inefficient in higher dimensions and

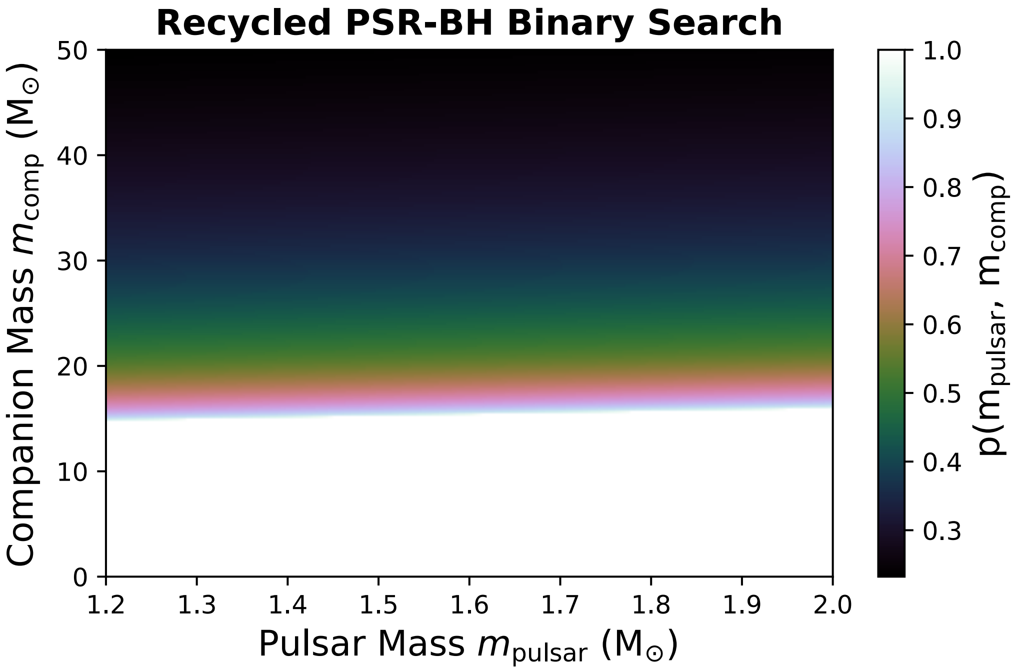


Figure 3.2: Fraction p of the total solid angle covered for our Recycled PSR-BH search. Region marked as 'white' is detectable for all inclination angles and orbital phases. For the rest of the diagram we retain only fractional sensitivity p where the orbital inclination angle is favorable.

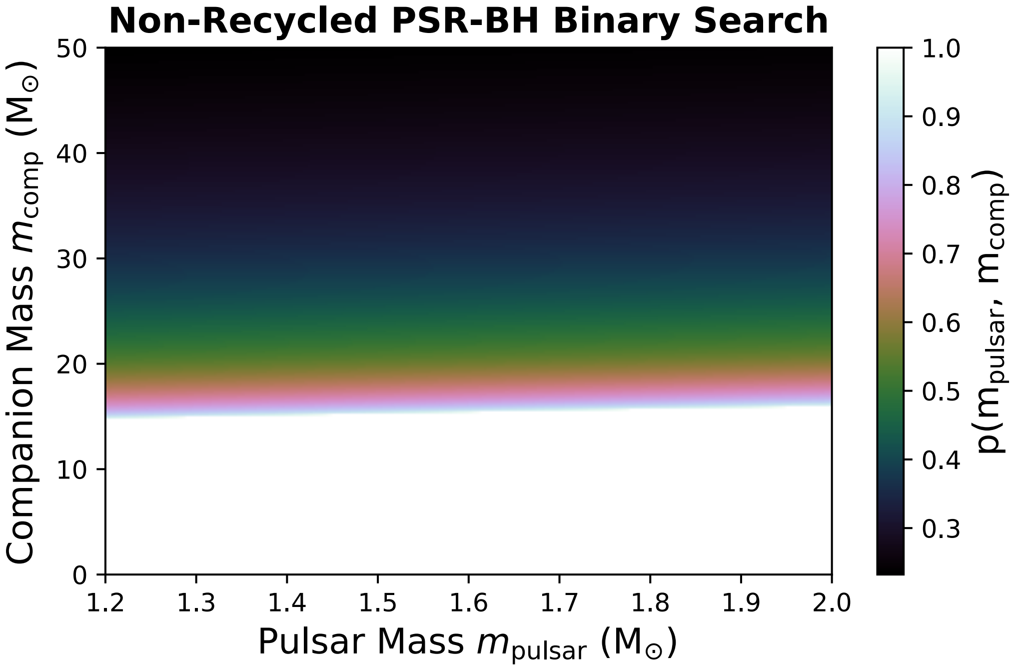


Figure 3.3: Fraction p of the total solid angle covered for our Unrecycled PSR-BH search. Region marked as 'white' is detectable for all inclination angles and orbital phases. Here we have a wider region where we are fully sensitive compared to our RPB search. However, the caveat is that this is true only for unrecycled pulsars with spin-period $P_{spin} \geq 100$ms and for binaries in circular orbits. For the rest of the diagram we retain only fractional sensitivity p where the orbital inclination angle is favorable.

for curved parameter spaces due to the difficulty in performing coordinate transformations to avoid non-constant metric components. Mismatch m as defined in equation 3.9 quantifies the fractional loss in detection statistic due to separation between the true location of the signal and the template compared to the ideal case when the template is exactly at the signal location. In practise, we defined a nominal mismatch value threshold i.e the worst-case tolerable detection statistic loss. In our searches, we fix this value to $m = 0.2$.

3.2.6 Random Template Banks

The main idea behind building random template banks is to give up the requirement of complete coverage of the parameter space ($\eta = 1$). Instead, we aim for a user-defined probabilistic coverage of the parameter space (η <1). We then distribute the random templates based on the chosen nominal value. For our searches, we chose $\eta = 0.9, m = 0.2$ similar to previous searches by Knispel (2011); Knispel et al. (2013); Allen et al. (2013). This implies that we cover 90 % of the parameter space with a nominal mismatch of 0.2. These values were chosen based on practical considerations depending on the total amount of required orbital templates for this configuration and the amount of computation resources available. The total required templates are distributed randomly with uniform probability based on values of the volume element which is the square root of the determinant of the parameter space metric $\sqrt{g}$ (see equation 3.10). The required number of templates can be calculated by computing the volume integral of this parameter space. Then we substitute values of the required coverage and mismatch in the template bank. We refer the reader to Messenger et al. (2009) for an in-depth review of this algorithm and section 5.2.3 of this thesis where we give the expressions required to calculate the number of orbital templates in a random template-bank for a given coverage and mismatch. For our RPB search, based on the orbital period, projected semi-major axis and initial orbital phase constraints given in section 3.2.4 and table 3.2, we need 16,157 orbital templates whereas for our NRPB search, we need 8389 orbital templates in our random template bank.

3.2.7 Distributing orbital templates in the parameter space using MCMC sampling

Once, we calculate the number of required orbital templates, we then need to distribute these templates optimally in the parameter space. Our goal here is to generate the template bank i.e sample from a probability distribution $P(\Omega, \tau, \alpha)$ which is proportional to the square root of the determinant of the metric tensor $\left|\sqrt{g}\right|$ which is our probability density function Q. Markov Chain Monte Carlo (MCMC) methods are well suited for such problems. Here, we specifically use the Metropolis-Hastings (MH) algorithm (Metropolis et al., 1953; Hastings, 1970). MH is useful to sample from a probability distribution $P(\Omega, \tau, \alpha)$ when direct sampling is cumbersome due to the difficulty involved in calculating the normalization factor. Here, we know that $P(\Omega, \tau, \alpha) = k\sqrt{g(\Omega, \tau, \alpha)}$ with k being the normalization factor. Briefly, MH works as follows- We start from a random point in the parameter x_n and then propose a new state x_{prop} and calculate the posterior distribution function (PDF) $\mathrm{P}(\mathrm{x}_{\text{prop}})$ at this new point. We always accept the step the PDF value increases and sometimes accept the step when the PDF value decreases. Here we implement a special case of MH called the Independence Chain Metropolis Hastings algorithm because our proposal density function Q does not depend on x_n i.e $\mathrm{Q}(x_{\text{prop}}|x_{\mathrm{n}}) = Q(x_{\text{prop}})$. Therefore, our acceptance criteria becomes

$$x_{\mathrm{n+1}} = \begin{cases} x_{\text{prop}}, & \text{if } \mathrm{U}(0,1) < \frac{\mathrm{P(x_{prop})}}{\mathrm{P(x_n)}} \\ x_{\mathrm{n}}, & \text{otherwise} \end{cases}$$

where U(0,1) is a uniform probability distribution between 0 and 1. We accept orbital templates using the above criteria until our required total templates are met. The scaling of orbital templates in the parameter space is non-uniform. Generally, more orbital templates are required for short orbital periods and a combinations of short orbital periods and high $a \sin \varepsilon$ values. We discuss extensively the details of this scaling relation and show a corner plot demonstrating the required sampling density across each Keplerian Parameter for PSR-BH binaries in circular and elliptical orbits in figures 5.1 and 5.2 respectively.

3.2.8 Stochastic Template Banks

As described earlier in equation 3.9, we use the parameter space metric in order to quantify the mismatch obtained if a binary pulsar signal is recovered at a different point (given by our orbital template) compared to its true location in the parameter space. However, the metric approximation is only valid up-to second order deviation of the coordinates. When higher-order terms of the metrics contribute, we typically overestimate the mismatch value i.e the metric returns a higher mismatch value than the true mismatch between the signal and the template. This leads to an overestimation of the required templates that can be pruned to save computation time without impacting the nominal mismatch in the template bank. This can be achieved using the stochastic template-bank algorithm (Harry et al., 2009). Here, we first start with a Random template-bank[3] and then we compute the mismatch between two templates at a time using the full integral (equation 3.7). If the two templates overlap each other by more than $1 - m$, then one of these templates is removed. We repeat this procedure for every pair until no new points are available. The drawback of this algorithm is that it is computationally expensive as for N templates in a template-bank can require up-to $\frac{N(N-1)}{2}$ evaluations of the integral in equation 3.7. This is an upper-limit as some templates maybe rejected quite early during pruning. In order for this approach to be practically useful, we need to parallelise the generation of this stochastic template-bank. This is done by dividing the template-bank into various subgroups and pruning overlapping templates within each subgroup in parallel. This is illustrated in figure

---

[3] Starting with a Random template bank is not a pre-requisite, however this leads to a faster convergence to create a stochastic template bank as the templates are already distributed based on the pseudo density function of the metric.

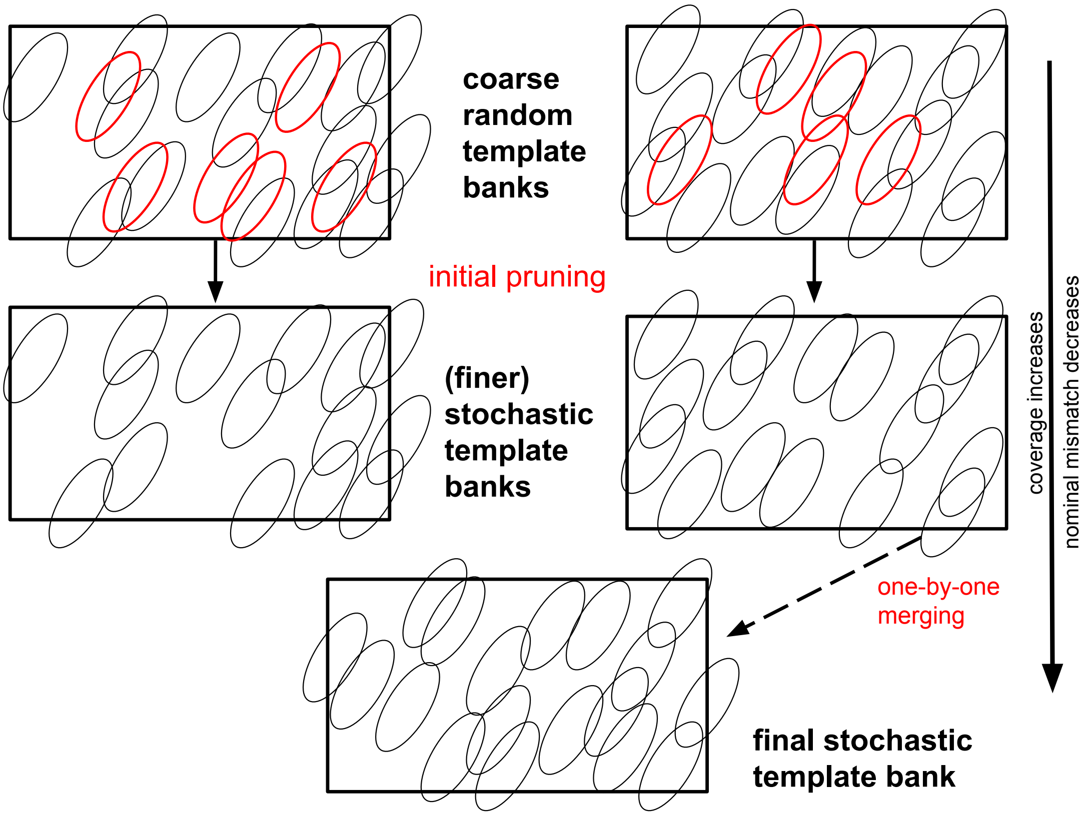


Figure 3.4: An illustration describing the pruning process in order to construct a stochastic template bank. We divide the random template-bank into various subdivisions (we show two subgroups here for simplicity) called coarse random template-banks. Templates that overlap by more than 1 - m, are pruned to form finer stochastic template banks. Two sub-groups can then be combined by one-by-one merging increasing our coverage and decreasing our nominal mismatch.

3.4. The main advantage here is we can start merging these subgroups in subsequent steps by comparing the mismatch between templates in one subgroup with another. Each template need not be checked against all the template of its own subgroup as this was already done in the previous step. The total number of iterations is still comparable to the serial case, however, since these are distributed across multiple CPU cores, large savings in wall-clock time is possible. For our RPB search, we had 16,157 orbital templates in our template-bank and after pruning we were left with 10,524 templates in the stochastic template-bank. This entire stochastic template bank generation process costed six days of computation time using 2000 CPU cores of the HERCULES Computing Cluster located in Garching. The total time taken for processing LOWLAT was expected to be about 18 months, therefore the up-front cost of investing a week worth of computation time in order to reduce the number of trials by 35 % was worth it in this case. For the NRPB search, we used a random template-bank for our pipeline as the total processing time savings wasn't significant.

3.2.9 Verification of Template Bank

Once we generate a random/stochastic template-bank it is important to verify that it satisfies the nominal coverage and mismatch values for signals located within its parameter space region. For brevity, we only show here the verification procedure used for our template-bank in the RPB search. A similar verification was also done for the template-bank created for the NRPB search. We do this by testing our template bank on simulated observations of PSR-BH binaries in circular orbits with a uniform distribution in orbital periods covering the 5-10 T_{obs} regime. We used a modified version of the code FAKE from SIGPROC[4] for these analysis. We simulated 10,000 observations with an integration time of 72-minutes with the same frequency bandwidth and number of channels compared to LOWLAT. The pulsar had a spin-period of 13 ms with a duty cycle of 20 %. We used a sampling time of 256 µs with 4-bits data rate since our main processing was done in this configuration. All our input parameters can be found in table 3.4. The maximum projected radius τ_{max} used depends on the orbital period, once these values were obtained we distributed these values with uniform probability between 0 and τ_{max}. Therefore, we sample all inclination angles between 0 and 90 degrees. Additionally, we did not place any constraints on the initial orbital phase α, they were allowed to vary uniformly between 0 and 2π radians. Our pulsar and companion masses were fixed for all observations at $1.4\mathrm{M}_{\odot}$ and $8.0\mathrm{M}_{\odot}$ respectively. In order to calculate mismatch (equation 3.9), we need to know the signal to noise ratio (S/N) for the "ideal" case assuming no losses from Doppler modulation. For this, we calculated the S/N recovered for an isolated pulsar with the same observational parameters given in table 3.4. Our results displaying the mismatch distribution is shown in figure 3.5. The median value of mismatch $\mathrm{m}_{0.5}$ of our template-bank is 0.19 i.e on average we lose 19% of S/N for binaries within the parameter space defined in table 3.4. The 90th percentile of the mismatch distribution $m_{0.9}$ is 0.31. This value is higher than our chosen nominal match value of 0.2 because of the extra mismatch that arises from searching in spin-frequency. Due, to the finite width of the FFT, signals can sometimes be located away from the centre of the fourier frequency bin, and in the worst case scenario between two bins which reduces our detection sensitivity. This effect is called *scalloping*. (See section 3.8.4 of Knispel 2011 or Ransom et al. 2002 for a detailed discussion on this). One technique to mitigate this to pad with data with zeros or mean value of the timeseries. We avoid this in our analysis due to our long integration time of 72-minutes in order to minimise computational time. An average extra 11 % loss due to a mismatch in frequency was acceptable for our searches. Based on these results, we concluded that the requirements of the template-bank have been fulfilled for our chosen search set-up.

---

[4] `https://github.com/SixByNine/sigproc`

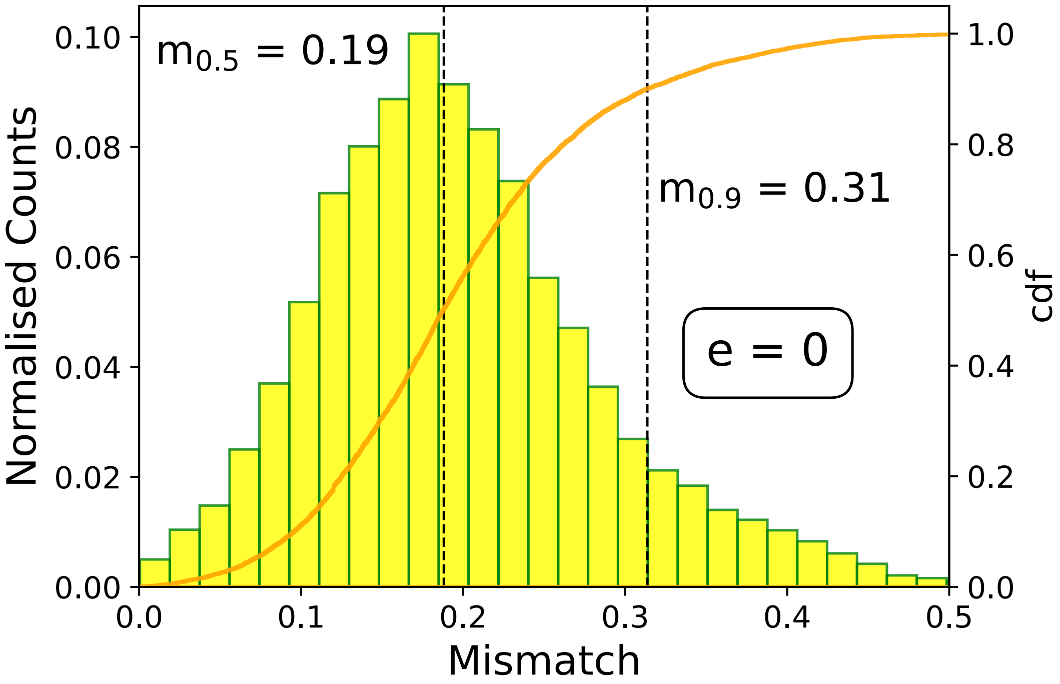


Figure 3.5: Histogram showing the mismatch distribution from the template-bank pipeline for simulated PSR-BH binaries in circular orbits. Vertical lines display the median and the 90th percentile of our mismatch distribution and the orange curve shows the cumulative distributive function.

Table 3.4: Input Parameters used for our Simulated Pulsar-Black hole binaries in circular orbits. We simulated 10,000 PSR-BH binaries with a Dispersion Measure DM = 0. Here, 'U' denotes a uniform distribution.

Parameter	Value/Range	Unit
t_{obs}	4320	s
$\mathrm{m}_{\text{pulsar,min}}$	1.4	$\mathrm{M}_{\odot}$
$\mathrm{m}_{\text{companion,max}}$	8.0	$\mathrm{M}_{\odot}$
Spin-Period	13.0	ms
Duty Cycle	20.0	%
Dispersion Measure	0.0	$\mathrm{pc\,cm^{-3}}$
Frequency Channels	1024	
Bandwidth	400	MHz
Frequency of Channel1	1181.804688	MHz
Channel bandwidth	0.390625	MHz
Signal-to-noise ratio of single pulse	0.02	
Data Bit-Rate	4	
Number of samples	2^{24}	samples
Sampling time	256e-06	s
Orbital Period (P_{orb})	U(6, 12)	h
Eccentricity (e)	0.0	
Initial Orbital Phase (α)	U(0, 2 π)	rad
Longtitude of Periastron (ψ)	0.0	rad
Projected Radius (τ)	$\mathrm{U}\left(0, \dfrac{G^{\frac{1}{3}}\Omega_{\text{orb}}^{-\frac{2}{3}} m_{\text{comp,max}}}{c\,(m_{\text{pulsar,min}} + m_{\text{comp,max}})^{\frac{2}{3}}}\right)$	lt-s

3.2.10 GPU Software

Using the template-bank algorithm can open up parameter space regions that are inaccessible using an acceleration or jerk search. See section 5.3.1 for a detailed comparison of acceleration search, jerk search and the template-bank pipeline for simulated PSR+BH binary observations and observations of the double pulsar PSR J0737-3039A. However, this comes at the cost of a much higher computation time. We require an order of magnitude more computational trials compared to the first-pass segmented acceleration search pipeline (Ng et al. 2015; Cameron et al. 2020) used on HTRU-S Lowlat. Using the same computing resources of previous searches i.e the HERCULES1[5] cluster operated by the Max Planck Computing and Data Facility (MPCDF) in Garching, Germany, the search component of the template-bank reprocessing pipeline for the entire Lowlat survey would have taken ∼ 8.7 years[6] and an additional ∼ 0.6 year (7 months) for folding, taking the total to ∼ 9.3 years. These numbers are conservative since they assume that the entire computation resource is dedicated solely to this project which isn't the case for a shared computing cluster. These timescales are unfeasible for the typical timescales of a PhD project and therefore, required the development of a faster approach. Graphical processing units (GPUs) are highly effective in speeding up the typical time-consuming parts of a pulsar-search pipeline including dedispersion, FFT and time-domain resampling. Therefore, this gave us a strong motivation to build a template-bank pipeline that can be run on GPUs. PEASOUP[7] is a GPU-Based Pulsar Search Pipeline built by Dr. Ewan Barr that implements a time-domain acceleration search (Johnston & Kulkarni, 1991). The main pipeline also implements dedispersion through the DEDISP library (Barsdell et al., 2012), Fourier de-reddenning, FFT, incoherent harmonic summing and a peak detection algorithm for candidate selection. Building on this work, I made a new version of this software called 3D-PEASOUP[8] that changed the time-domain resampling and candidate reduction algorithm of the pipeline. The orbital templates for a particular search are computed offline using CPU scripts and is passed on-to this software as an input file. We go into more details about the time-domain resampling algorithm for the general case of elliptical orbits in section 5.5.1. Bright pulsars, can be detected with multiple orbital templates with varying S/N or candidate significance. Therefore, we need to group these 'similar' candidates together, keeping only the candidate with the highest S/N in order to save computation time during folding. We used the standard top-down consolidation algorithm (Knispel, 2011), to reduce the number of duplicate detections. Variants of this algorithm with the formula adapted for polynomial searches have also been previously implemented

---

[5] This cluster was recently upgraded in March 2021 and is now called HERCULES2 with more GPUs installed which substantially saves computation time for pulsar searches. However, since we commenced our searches in 2019, we assume numbers from HERCULES1.

[6] This number is calculated assuming a DM smearing tolerance of 10 % which gives 1876 DM trials between 0 and 3000 $\mathrm{pc\,cm^{-3}}$. Lowlat survey consists of 1230 pointings, each consisting of 13 beams, and the average bench-marked computation time for each trial is 4.38 seconds for a 72-minute observation with a sampling time of 256 micro second. Therefore, total processing time = DM trials × total orbital templates × Average Time for 1 trial × pointings × beams/(Total CPU Cores = 5000.)

[7] `https://github.com/ewanbarr/peasoup`

[8] `https://github.com/vishnubk/3D_peasoup`

in time-domain search softwares like SIGPROC[9] and PEASOUP. Lowlat was reprocessed on the upgraded OzSTAR[10] High Performance Cluster (HPC) located in Swinburne University of Technology, Australia. OzSTAR has a 5 PB Lustre-zfs filesystem with 4140 SkyLake CPU cores clocking at 2.3GHz across 107 standard compute and eight data crunching nodes along with 230 NVIDIA Tesla P100 12 GB GPUs (2 per node). We refer the readers to the online documentation which gives the full list of the cluster specifications. We performed detailed bench-marks on this computing cluster comparing the CPU and GPU pipeline in order to quantify our speed-up factor.[11] The CPU based pipeline performed dedispersion using the DEDISPERSE_ALL routine in SIGPROC[12] followed by a PYTHON-based implementation of the nearest neighbour time-domain resampler (section 5.5.1) that was accelerated using NUMBA[13] which is a just-in-time (JIT) compiler that translates python code into machine-code. Folding for both pipelines is done a CPU using the PREPFOLD program from PRESTO[14]. We parallelized the single core programs like our resampler and PREPFOLD by using the MULTI-PROCESSING library in PYTHON. Our benchmarks comparing both these pipelines can be found in table 3.5. We compare the single-core performance between a CPU and a GPU as well as the total folding time which is the same for both pipelines. Overall, we achieve a speed-up factor of 400 for the search component[15] compared to the CPU pipeline benchmarked on OzSTAR and a factor of $\sim$ 700 compared to CPUs in HERCULES1. The interesting aspect to note here is that we are currently limited by folding time whereas traditionally pulsar searches were limited by the time required to FFT different timeseries from each binary search trial. The overall net improvement factor taking into account folding time is 16.5 which is still significant! Therefore, the development and subsequent application of the GPU pipeline were crucial in enabling us to reprocess Lowlat within a feasible amount of time. Finally, reprocessing Lowlat typically produce tens of millions of pulsar candidates, therefore manual inspection of these candidate to distinguish between a pulsar and radio frequency inteference/noise is not feasible. We will expand on a novel semi-supervised machine learning pipeline developed during this PhD in chapter 4. This pipeline also played a crucial role in helping us find new pulsars within Lowlat.

---

[9] The program BEST in sigproc applies the candidate reduction algorithm.

[10] `https://supercomputing.swin.edu.au/ozstar/`

[11] The astute reader would have noticed that the CPU benchmarks given for HERCULES1 and the CPU-benchmark values in table 3.5 are different. This is because, values given in table 3.5 are benchmarked on OZSTAR which has more modern CPUs and the single-core performance is approximately a factor of 2 better than a CPU core in HERCULES1.

[12] `http://sigproc.sourceforge.net/`

[13] `https://github.com/numba/numba`

[14] `https://github.com/scottransom/presto`

[15] includes dedispersion, Fourier de-reddenning, FFT, incoherent harmonic summing and a peak detection algorithm for candidate selection

Table 3.5: Computational time comparison between the CPU and GPU based Pipeline processsing a single 72-minute Lowlat observation. These benchmarks have been run on 10 % of our typical search range and total candidates folded i.e 189 DM trials, 1000 template-bank trials per DM and folding 200 candidates per beam.

	Operation	CPU Based Pipeline Single-Core time (mins)	Template-Bank Peasoup 1-GPU (mins)	Speed up Factor
Periodicity Search	Dedispersion	56.4	0.37 ($\approx$ 22.1s)	152.4
	Search	8332.4 ($\approx$ 138.9 hrs)	20.7	402.5
	Search Time Subtotal	**8388.9**	**21.0**	**$\sim$ 400**
Candidate Production	**Folding Time Subtotal**	**520.1**	**520.1**	1
	Total Time: Single Core	**8909**	**541.1**	**16.5**

3.3 Results

3.3.1 New Pulsar Discoveries

A total of **20** new pulsars have been discovered and confirmed in the reprocessing of Lowlat using the template-bank pipeline including a new millisecond pulsar which has been confirmed to be in a binary orbit. A further three ClassA candidates (defined as highly likely to be real) await confirmation. Additionally, we re-detected **50** new pulsars that were missed by the first-pass processing of Lowlat (Ng et al., 2015; Cameron et al., 2020). Out of these, 47 pulsars were discovered in the reprocessing of 72-minute observations of Lowlat using PEASOUP with a combination of higher acceleration range $|a_{\rm max}| = 100\,{\rm m\,s^{-2}}$ on the decimated Lowlat data - $t_{\rm samp} = 256\,\mu{\rm s}$ and a mild acceleration search i.e $|a_{\rm max}| = 2\,{\rm m\,s^{-2}}$ on the full time resolution of the data - $t_{\rm samp} = 64\,\mu{\rm s}$ (Sengar et al., prep). An additional 3 pulsars were discovered by applying a Fast-Folding Algorithm pipeline (Wongphechauxsorn et al., prep). Given the large number of re-processing discoveries from Lowlat in a short period of time due to searches been done on GPUs, compared to our first pass search which found 100 new pulsars over the time span of 6 years, a full 1-year timing followup on all pulsars have not yet been possible. Traditionally, once pulsars have been identified in surveys like HTRU and PMPS, we typically need to do gridding observations around the position of the discovery beam. One such strategy is by placing the beam in the 'N-E-S-W' direction along with an observation in the discovery beam in order to get a better position of the pulsar (Morris et al., 2002). For pulsars in lowlat, this would require a total of 1.2 hrs $\times$ 5 = 6 hours per pulsar. Applying the 'Ring-of-3' grid described in Ng (2014) would also require 3.6 hours per pulsar which is not feasible given the large number of discoveries that needed to be monitored and the extremely competitive Galactic centre time at Parkes. Therefore, follow-up observations beyond confirmation have currently been restricted mostly to millisecond and binary pulsars. Few pulsars have extremely well known positions with a position uncertainty of $\sim 30''$ due to these pulsars being redetected in the ongoing L-Band MPIfR Galactic Plane survey (MGPS-L, Padmanabh et al. prep) survey using the MEERKAT Radio telescope. Once promising candidates were identified, we performed search-mode follow-up observations using the ultra-wide-bandwidth, low-frequency receiver (UWL, Hobbs et al. 2020) installed at the Parkes radio telescope to confirm these pulsars. A full list of the new pulsars described here along with their position, spin period, DM, and DM dependent distance based on the 'NE2001' model (Cordes & Lazio, 2002) and the YMW16 model (Yao et al., 2017) can be found in table 3.6. The integrated pulse profiles from the discovery observations of these new pulsars can be found in figure 3.6.

Table 3.6: Discovery parameters of the 20 newly-discovered pulsars for which full timing solutions are not yet available. Most of these pulsars have been assigned a temporary name with only two digits of declination since our position uncertainty is at least 14′ - the FWHM of the Parkes beam . Few pulsars have extremely good position measurements (uncertainty: $\sim 30''$) due to them being redetected in the ongoing MGPS-L survey using MEERKAT. The reported barycentric spin period P, dispersion measure DM and barycentric period epoch PEPOCH values have been taken from the confirmation observation, and values in parenthesis represent our uncertainty in the final digit. We also report DM-dependent distances based on 'NE2001' model (Cordes & Lazio, 2002) and the YMW16 model (Yao et al., 2017)

PSRJ	pointing/beam	RA (J2000) (hms)	Dec (J2000) (dms)	gl (°)	gb (°)	P (ms)	DM ($\mathrm{pc\,cm^{-3}}$)	PEPOCH MJD	S/N	D_{YMW16} (kpc)	D_{NE2001} (kpc)
J1142−62	2011-05-19-08:08:26/01	11:42:39.24	-62:21:52.2	295.050	-0.550	1573.49(1)	267(33)	59015.204613	16.42	2.49	5.21
J1251−59	2011-04-20-15:54:35/09	12:51:13.0	-59:57:36.4	302.904	2.912	1066.630(6)	207(7)	58995.340754	12.84	5.25	4.37
J1359−59	2011-12-21-16:53:53/08	13:59:12.12	-59:17:09.03	311.499	2.454	374.1318(4)	323(4)	59099.151304	13.18	6.63	6.50
J1538−57	2010-12-11-23:55:11/05	15:38:01.32	-57:16:33.2	323.992	-1.380	205.61668(8)	543(2)	59171.893788	15.23	8.31	8.55
J1539−59	2012-08-05-12:04:48/10	15:39:00.72	-59:57:27.16	322.503	-3.615	189.0307(2)	209(4)	59171.955067	12.17	5.44	4.18
J1558−5403[a]	2012-08-05-03:42:42/10	15:58:30.93	-54:03:08.6	328.302	-0.724	583.705(2)	783(18)	59048.325298	19.14	7.66	10.21
J1609−5225[a]	2012-08-02-09:33:26/09	16:09:02.41	-52:25:18.8	330.516	-0.468	231.8080(7)	993(10)	58936.863771	16.54	7.20	14.32
J1620−47	2011-12-22-00:16:31/06	16:20:14.28	-47:44:29.5	335.045	1.681	386.8123(4)	396(4)	59206.817811	13.06	6.49	6.17
J1629−44	2011-07-04-08:36:43/04	16:29:02.80	-44:08:17.1	338.693	3.110	365.0362(3)	310(3)	59212.949630	12.39	11.68	5.64
J1637−44	2011-07-03-10:59:47/02	16:37:27.84	-44:17:18.30	339.608	1.901	274.1809(2)	271(3)	59205.876584	11.96	5.31	4.20
J1657−46	2012-11-25-23:53:23/13	16:57:49.66	-46:24:39.4	340.300	-2.200	892.383(4)	652(6)	59370.741151	10.53	23.74	13.16
J1718−37	2011-05-17-14:17:08/10	17:18:24.48	-37:27:18.0	349.762	0.096	591.629(1)	693(7)	59113.227057	9.90	4.74	7.64
J1722−38	2011-06-27-14:48:40/04	17:22:28.20	-38:09:57.6	349.637	-0.969	523.782(2)	679(15)	58995.500900	14.39	3.40	3.29
J1723−37	2011-05-19-15:32:31/04	17:23:12.48	-37:57:23.9	349.891	-0.970	150.92019(7)	266(2)	59206.190774	11.49	3.74	3.82
J1737−33	2010-12-12-03:46:51/05	17:37:40.80	-33:05:20.69	355.575	-0.750	634.00(1)	497(59)	59048.428250	14.38	4.51	5.80
J1743−24	2012-11-28-05:31:51/11	17:43:43.50	-24:27:25.90	3.600	2.700	11.757463(2)	196.8(6)	59370.647086	14.04	4.44	3.60
J1754−25	2012-07-26-12:48:29/01	17:54:38.45	-25:52:48.9	3.650	-0.150	256.8954(1)	814(3)	59206.969001	10.65	4.98	10.93
J1806−17	2011-07-05-13:09:31/07	18:06:31.68	-17:21:46.25	12.415	1.680	280.7599(4)	539(5)	59077.385176	13.35	18.99	8.73
J1806−19[b]	2012-12-11-02:16:06/05	18:06:49.44	-19:26:25.8	9.700	0.300	1101.568(7)	558(14)	50682.454323	10.50	4.51	6.77
J1807−23	2012-08-05-14:33:10/13	18:07:15.6	-23:13:18.9	7.381	-1.324	890.195(2)	295(3)	59187.126936	17.31	4.30	4.84

[a] Positions reported for these pulsars were obtained from the ongoing MGPS-L survey (Padmanabh et al., prep) but the Period, DM, PEPOCH and S/N reported in these rows are from a follow-up Parkes Observation.

[b] This pulsar was confirmed in an ongoing reprocessing of PMPS observations (Sengar et al., prep).

3.3.2 PSR J1753−2819 - A rare intermediate spin-period pulsar in orbit around a light companion

PSR J1753−2819 is a millisecond pulsar with a spin-period of 18.62 ms that was detected during the re-processing of Lowlat. PSR J1753−2819 was initially announced in Mickaliger (2013) and later expanded on in Perera et al. (2019). However, this pulsar was not included in the pulsar catalog PSRCAT[16] or the Galactic MSP catalog[17] around the time we initially found this pulsar. Therefore, this pulsar was presumed to be a LOWLAT discovery until recently. We first present an independently derived orbital solution of PSR J1753−2819 from observations taken at the Parkes Radio telescope between December 2020 and July 2021. We then present a combined and updated timing solution including archival and recently obtained TOAs from Jodrell Bank Observatory (JBO) along with the TOAs collected from Parkes. TOAs from JBO were provided to us by B.Stappers (2021, priv. comm.). PSR J1753−2819 has the eleventh highest dispersion measure of 298.4 pc/cm^3 out of 403 currently known Galactic field MSPs[18] not associated with a Globular Cluster. Our initial redetection plot from LOWLAT is shown in figure 3.7. The barycentric spin-period of this pulsar in the LOWLAT observation was $P_{\text{spin}} = 18.623\,025(3)$ ms with the number in the parenthesis describing our uncertainty in the last decimal digit. We independently confirmed this pulsar with a search-mode observation taken at the Parkes Radio telescope where we detected a barycentric spin-period of $P_{\text{spin}} = 18.622\,979(2)$ ms. The apparent spin-up of this pulsar hinted that the pulsar could be in a binary orbit. Following this, we regularly monitored this pulsar between December 2020 and July 2021 with at least a monthly cadence with the PARKES radio telescope. We used the UWL receiver in conjunction with the MEDUSA backend in search-mode in order to constrain the orbit of the pulsar and collect more TOAs. Our observations were recorded covering the full spectral range offered by UWL between 704 MHz and 4032 MHz. For our analysis, we restrict ourselves to a narrow band between 1100 - 1800 MHz which has approximately twice the bandwidth of the original LOWLAT observation taken with the MB20 receiver. The influence of RFI in the lower frequency sub-band hindered our ability to detect this pulsar and extract TOAs. We did not detect any pulsations from this pulsar in the higher frequency sub-bands as well and we leave detailed investigations of the properties of this pulsar in high-frequency observations for future studies. Follow-up observations undertaken as part of the ongoing L-Band MPIfR Galactic Plane survey (MGPS-L, Padmanabh et al. prep) using the MEERKAT telescope on 16th June 2021 (MJD: 59381) gave us a improved sky position of the pulsar RA (J2000) = $17^{\text{h}}53^{\text{m}}55.39^{\text{s}}$, Dec(J2000)= $-28^\circ 19^{\text{m}} 42.9^{\text{s}}$ with our position uncertainty in the order of $30''$. This greatly enhanced our detection sensitivity and future follow-up time required at Parkes along with getting a phase-connected timing solution.

---

[16] This has since been fixed in version 1.65 of the catalog released on 29/07/2021.

[17] `http://astro.phys.wvu.edu/GalacticMSPs/GalacticMSPs.txt`

[18] See footnote 19

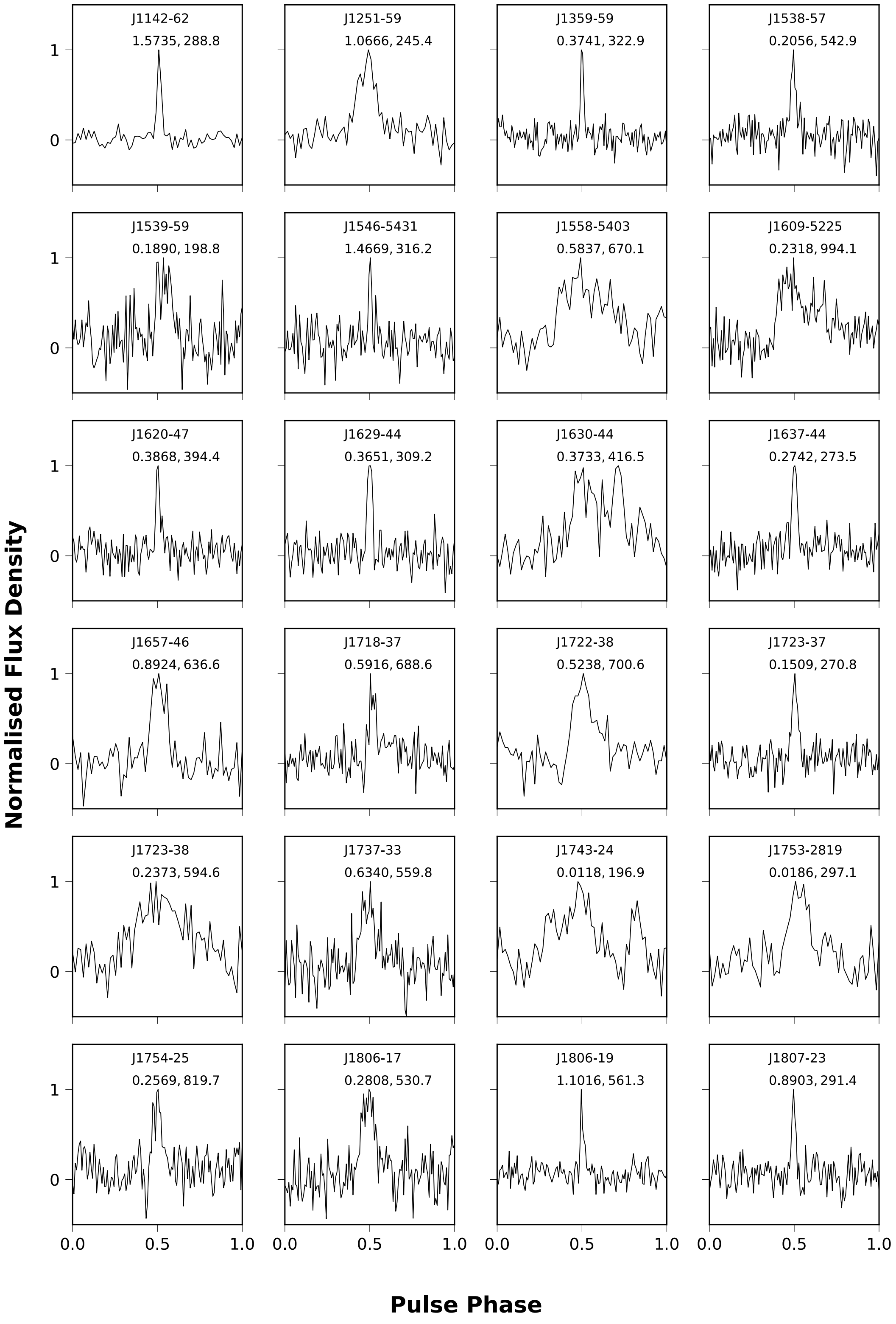


Figure 3.6: Integrated Pulse profile of the 20 new discoveries, PSR J1753−2819 and 3 ClassA candidates found in the reprocessing of LOWLAT. Each pulse profile consists of 64 phase bins. We show here normalised flux density and the pulse profile has been rotated, to peak at 0.5 rotational phase. At the top, we indicate the current pulsar name, spin-period in seconds and DM in pc cm^{-3}.

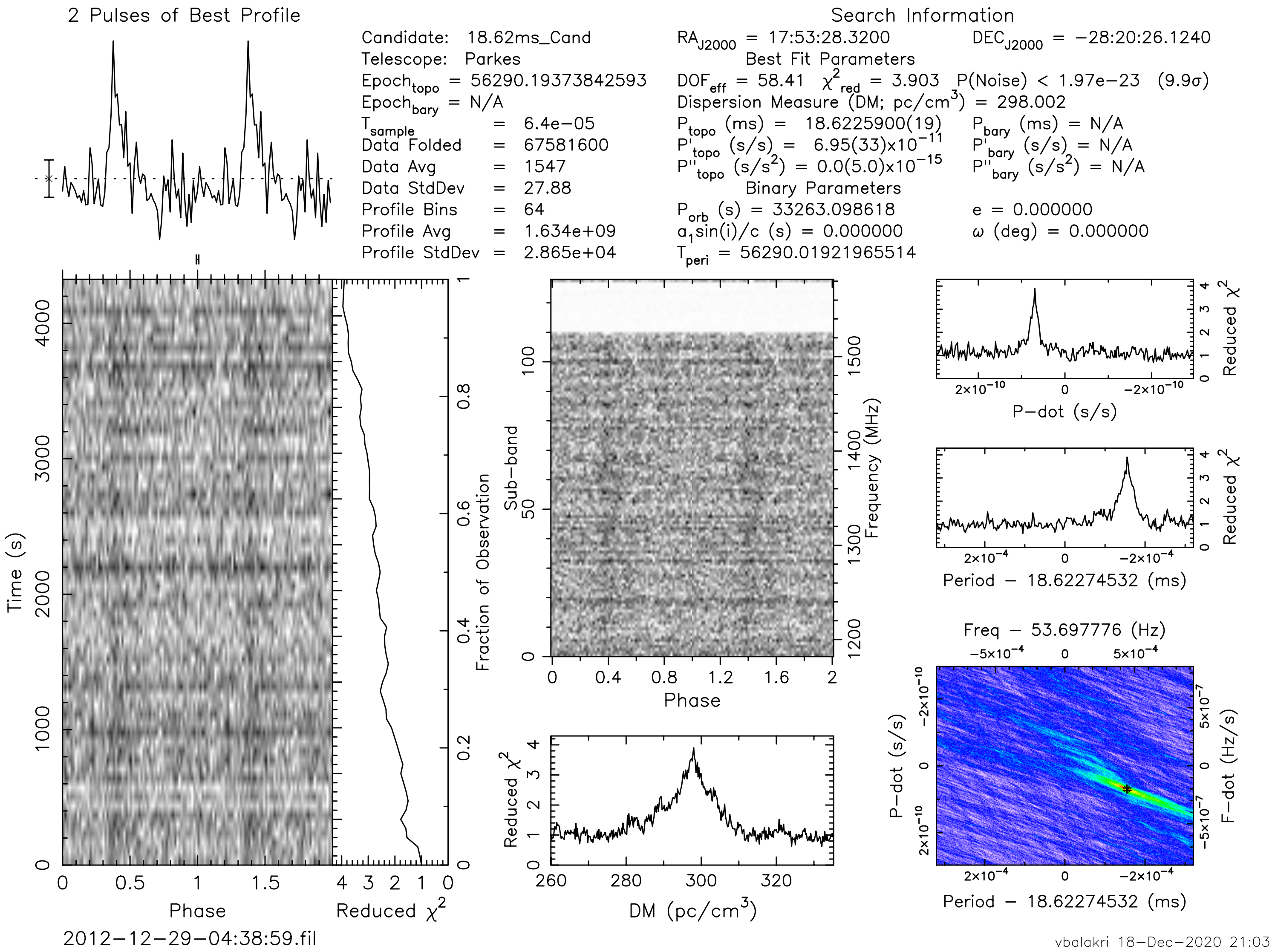


Figure 3.7: Initial Pulsar Detection plot of PSR J1753−2819 found in the LOWLAT beam 2012-12-29-04:38:59/11. The orbital template that was used to detect the pulsar can be found in the top right under 'Binary Parameters'.

3.3.2.1 Preliminary estimates of orbital parameters

As the pulsar moves around its orbit, the observed apparent barycentric spin-period of the pulsar changes due to Doppler modulation. Therefore, it is important to determine the Keplerian orbital parameters of the system in order to get a coherent timing solution and for further astrophysical analysis. In this section, we describe the techniques used to arrive at the orbital parameters of the pulsar PSR J1753−2819. Two readily available measurements from a search-mode pulsar observation is the spin-period P and the spin-period derivative $\dot{P}$. We closely follow the technique described in Freire et al. (2001) where we plot the apparent spin-period and line of sight acceleration measurements of the pulsar from each observation. The advantage of this technique is that in the period-acceleration plane, we get a parametric curve that does not depend on time and hence does not require solving Kepler's equation. Acceleration can be calculated from P and $\dot{P}$ using the relation

$$a = \frac{c\dot{P}}{P}. \tag{3.20}$$

For circular orbits, the apparent spin-period of the pulsar from Roy (1988); Freire et al. (2001) can be written as:

$$P(\theta) = P_0\left(1 + x\frac{2\pi}{P_B}\cos\theta\right) = P_0 + P_1\cos\theta, \tag{3.21}$$

and

$$A(\theta) = \frac{-4\pi^2}{P_B^2}x\sin\theta = -A_1\sin\theta, \tag{3.22}$$

where $P_1 = P_0 x\frac{2\pi}{P_B}$ and $A_1 = \frac{4\pi^2}{P_B^2}$, x is the projected semi major axis in light seconds, P_B is the orbital period of the pulsar, P_0 is the intrinsic barycentric rotational period of the pulsar and θ is the orbital phase.

It is clear from the right hand side of equations 3.21 and 3.22 that these are parametric equations of an ellipse centered at $(P_0, 0)$ with the length of the horizontal and vertical semi-axes being P_1 and A_1 respectively. Finding the best fit ellipse can be quit difficult from a computational point of view. Following the recommendation given in Appendix A of Freire et al. (2001), squaring each of the equations 3.21, 3.22 and multiplying one of the squares with a constant and summing them, we get the parametric equation of a parabola:

$$A^2 = a_2P^2 + a_1P + a_0. \tag{3.23}$$

We apply a linear least square fit in the (A^2, P) plane to determine the coefficients a_0, a_1 and a_2 respectively. Following this the orbital parameters can be calculated as:

$$P_0 = \frac{-a_1}{2a_2}, \tag{3.24}$$

$$P_B = \frac{2\pi c}{P_0\sqrt{-a_2}}, \tag{3.25}$$

$$x = \frac{P_B}{2\pi P_0}\sqrt{P_0^2 - \frac{a_0}{a_2}}. \tag{3.26}$$

For each observation, we take the spin-period and spin-period derivative values including their uncertainties returned from the '.bestprof' files in PRESTO. These measurements along with our best-fit model can be found in figure 3.8. Using this we get $P_0 = 0.0186236(3)$ s, $P_B = 0.3899(6)$ days and $x = 0.2090(3)$ lt-s.

3.3.2.2 Refining Orbital Period Measurement using Lomb-Scargle Periodogram

Lomb-Scargle Periodogram (Lomb, 1976; Scargle, 1982) is a commonly used algorithm to find periodicity of signals in unevenly sampled data. This is applicable for our case since our follow-up observations depends on the allocation of telescope time, weather conditions and appropriate sidereal time, therefore we unevenly sample the orbit. We apply this algorithm on data with multiple period measurements in the form of a timeseries. We then extract the period that corresponds to the maximum value in the Fourier power spectrum to detect the orbital period of the binary pulsar. The drawback of this algorithm is that it typically requires many observations in order to get a reliable measurement of the orbital period. Therefore, we applied this technique as a second step after collecting measurements from several follow-up observations at Parkes. Using this technique as shown in Figure 3.9, we were able to get a better measurement of the orbital period of the pulsar $P_B = 0.3876(7)$days.

3.3.2.3 Conventional Period Analysis using FITORBIT

We can further refine our estimates for all Keplerian parameters by conventional period analysis. The two techniques described above give us only an approximate measurement of the orbital period, and projected semi major axis of the orbit. This can be confirmed by plotting the difference between the observed period and the period obtained from our binary model predictions. We can use these differences to refine our estimates for each of our Keplerian orbital parameters using the software FITORBIT[19] originally written by A. Brinklow and A. Lyne. The software tries to minimise the sum of the squares differences of the intrinsic orbital period and the orbital period measurements predicted by the pulsar ephemeris for the same epoch at the barycenter by varying other ephermeris parameters chosen by the user. This is implemented using the Levenberg-Marquardt Method (LMM; Press et al. 1992). We refer the interested readers to section 4.4 and appendix A.3 of Freire (2000) which goes into more technical details on how each of the orbital parameters is fine tuned in FITORBIT.

3.3.2.4 Timing Solution of PSR J1753−2819 using TEMPO/TEMPO2

Following this, we determined a phase-connected timing solution for PSR J1753−2819 using the procedure described earlier in section 1.7. We used DSPSR[20] (van Straten

---

[19] `https://github.com/vivekvenkris/fitorbit`

[20] `http://dspsr.sourceforge.net/`

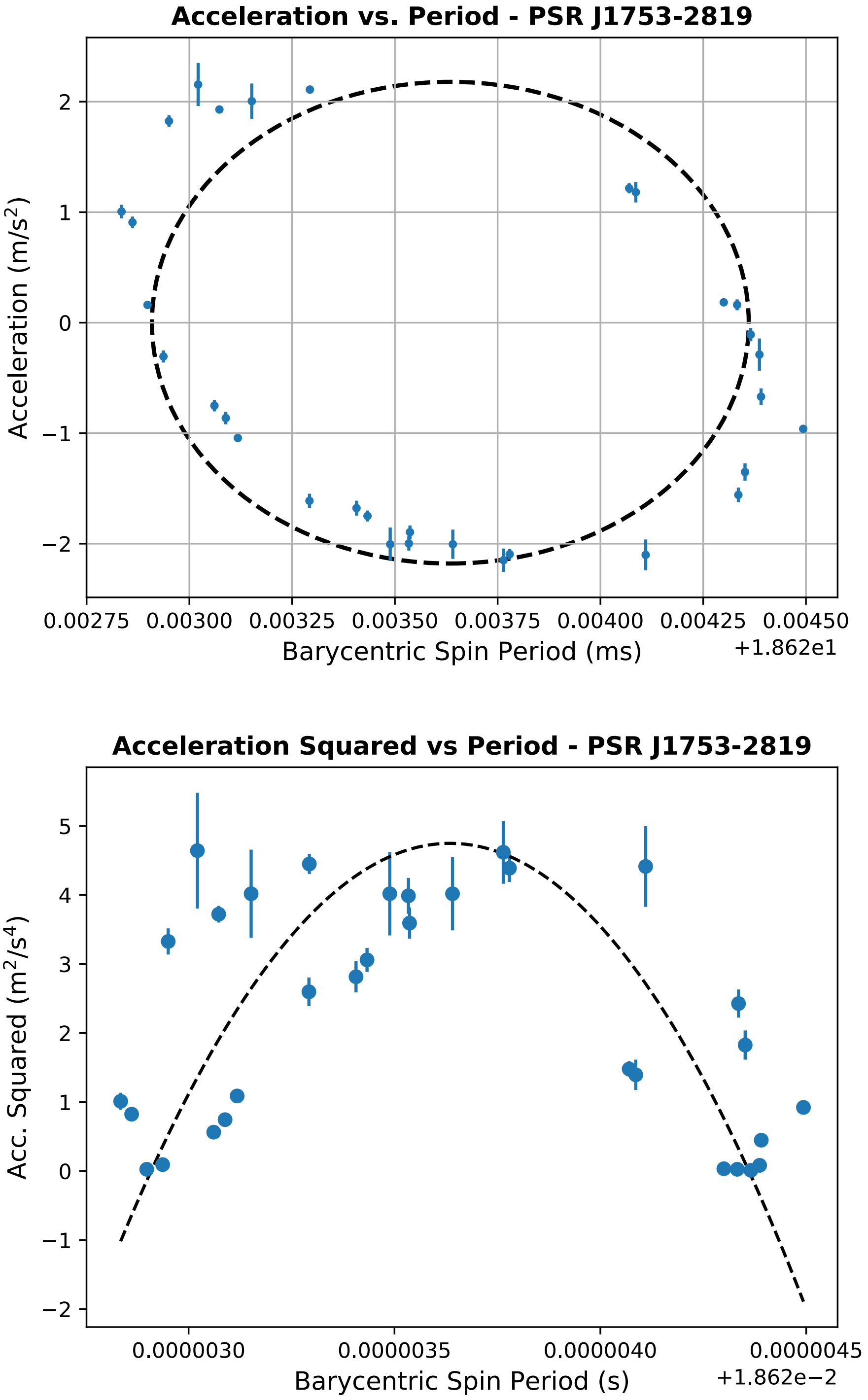


Figure 3.8: Observed barycentric acceleration and spin-period plotted in a period-acceleration plane (top) and period-acceleration square plane (bottom). The error bars indicate 1-sigma errors in period and acceleration. Black dashed line indicate the best-fit model which was used to calculate the Keplerian orbital parameters of this pulsar. A caveat to be added here is that the period and acceleration uncertainties displayed in these figures are based on the values returned by PREPFOLD which is a folding routine part of the PRESTO software package. These uncertainties are almost certainly an underestimate of the true uncertainty and as such any orbital fit values derived from these plots are only an approximate solution of the orbit of the pulsar.

& Bailes, 2011) to dedisperse and fold our data with the best period and acceleration value detected in the FFT. We then obtained TOAs by cross-correlated the resulting pulse profile with a noise-free template derived from the best detection of this pulsar using routines in PSRCHIVE[21] (van Straten et al., 2012). We used the timing software TEMPO[22] to obtain our solution following closely the phase jump technique described in Freire & Ridolfi (2018). Following this, we combined our data with updated TOAs from JBO and we present here an updated timing solution of PSR J1753−2819. The most commonly used binary model in pulsar timing is the Blandford & Teukolsky (BT,Blandford & Teukolsky 1976). However, using the BT model for pulsar orbits with low values of eccentricity as is the case for PSR J1753−2819, produces highly covariant values between the epoch and longitude of periastron. To avoid this, Lange et al. (2001) introduced a binary timing model called ELL1 where Epoch of periastron (T0) and eccentricity e is replaced by the epoch of Epoch of ascending node (TASC) and two Laplace-Lagrange parameters (η and κ). Using our best fit parameters of TASC, η and κ, we can infer the T0 and orbital eccentricity of the pulsar orbit. We use this ELL1 model to derive our timing solution and our best-fit timing parameters for PSR J1753−2819 can be found in table 3.7. We report here an updated measurement of the period derivative of this pulsar $\dot{P} = -4.967(7)\times10^{-20}\,\mathrm{ss}^{-1}$ consistent with the values reported earlier in Perera et al. (2019). This negative value of $\dot{P}$ is unusual as explained earlier in section 1.4 pulsars are expected to have an intrinsic positive period derivative due to loss of rotational kinetic energy. Perera et al. (2019) reported that this negative observed $\dot{P}$ is not intrinsic and likely caused by the acceleration of this pulsar along its lines of sight due to the Galactic gravitational field. Using the Galactic model presented in Sofue (2013), they reported a model-dependent intrinsic $\dot{P}_{\mathrm{int}} < 2.08 \times 10^{-19}\,\mathrm{ss}^{-1}$. We refer the interested reader to their paper which goes into more details about how this value was calculated.

In the next section, I will first describe one of our new binary pulsar discoveries PSR J1743−24 which is another unusual intermediate spin-period pulsar similar to PSR J1753−2819 but in a much longer orbit. I will then present a combined discussion on both these pulsars by focusing on its unusual combination of spin period and companion mass properties by putting it into context of other known binary pulsars and its likely that the formation of both these pulsars is inconsistent with standard binary stellar evolution models.

---

[21] http://psrchive.sourceforge.net/

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

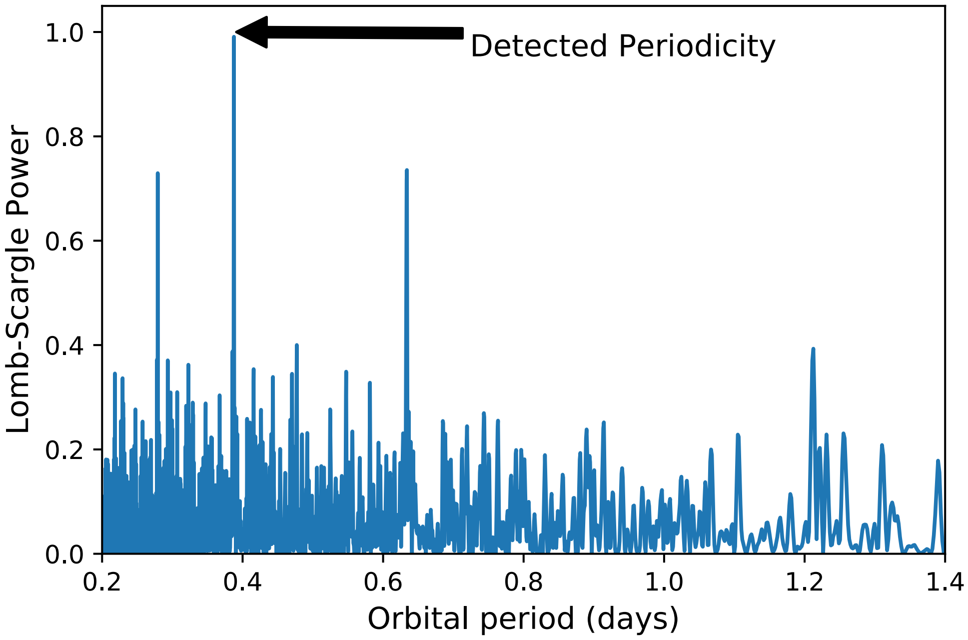


Figure 3.9: A Lomb-Scargle Periodogram constructed from our measurements of spin-period at different observing epochs. The detected periodicity (orbital period) is given by the peak value of the Lomb-Scargle Power.

Table 3.7: Best fitting timing parameters of PSR J1753−2819 obtained from fitting updated TOAs from JBO and Parkes with TEMPO2. Values in parenthesis represent 1 σ uncertainties on the final digit. The time units are Barycentric Coordinate Time (TCB) and the Solar system ephemeris used is JPL DE436.

Pulsar	J1753−2819
Right Ascension, α (J2000)	17:53:56.184(5)
Declination, δ (J2000)	−28:19:29.96(6)
Spin Frequency, ν (s^{-1})	53.695219546226(10)
1st Spin Frequency derivative, $\dot{\nu}$ (Hz s^{-2})	$1.432(2)\times10^{-16}$
Reference Epoch (MJD)	57299
Start of Timing Data (MJD)	56173.742803402153832
End of Timing Data (MJD)	59434.894515986230434
Dispersion Measure, DM (pc cm^{-3})	298
Solar System Ephemeris	DE436
Terrestrial Time Standard	TT(BIPM2015)
Time Units	TCB
Number of TOAs	207
Residuals RMS (μs)	129.639
Binary Parameters	
Binary Model	ELL1
Projected Semi-major Axis, $x_{\rm p}$ (lt-s)	0.20967(1)
First Laplace-Lagrange parameter, η	$-1(1)\times10^{-4}$
First Laplace-Lagrange parameter, κ	$1(1)\times10^{-4}$
Epoch of passage at Ascending Node, $T_{\rm asc}$ (MJD)	59206.093761(6)
Orbital Period, P_b (days)	0.387677374(1)
Inferred Orbital eccentricity, e[a]	0.0002(1)
Derived Parameters	
Spin Period, P (s)	$1.8623631832609(3)\times10^{-2}$
1st Spin Period derivative, $\dot{P}$ (s s^{-1})	$-4.967(7)\times10^{-20}$
Mass Function, $f(M_{\rm p})$ ($M_\odot$)	$6.585(8)\times10^{-5}$
Minimum companion mass, $M_{\rm c,min}$[b]($M_\odot$)	5.17×10^{-2}
Median companion mass, $M_{\rm c,med}$[c]($M_\odot$)	6.00×10^{-2}

[a] Derived from ELL1 binary model.

[b] $M_{\rm c,min}$ is calculated assuming the orbital inclination angle $i = 90$ and pulsar mass M_p is 1.4 $M_\odot$.

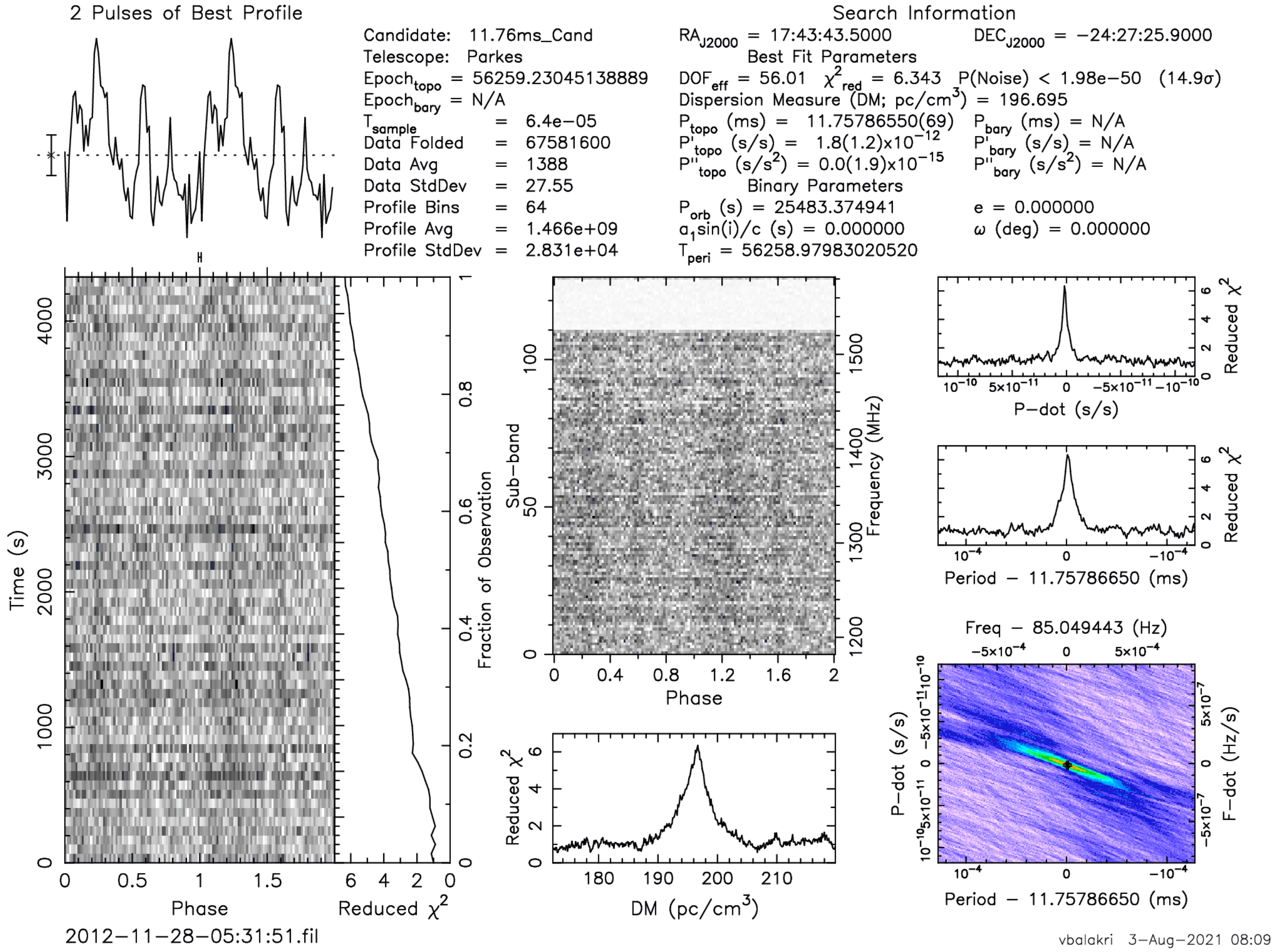
[c] $M_{\rm c,med}$ is calculated assuming the orbital inclination angle $i = 60$ and pulsar mass M_p is 1.4 $M_\odot$.

3.3.3 PSR J1743−24, another rare intermediate spin-period binary pulsar in a 70.7 day orbit

PSR J1743−24 is a millisecond pulsar discovered during our reprocessing with a barycentric spin period of 11.757 463(2) ms and a detected DM of 196.7 pc cm^{-3}. Our discovery plot can be found in figure 3.10. This pulsar has since been confirmed with follow-up observations taken at the Parkes Radio telescope using the same configuration as described in section 3.3.2. PSR J1753−24 has a relatively wide pulse profile with a duty cycle of 29.7%. The shape of the integrated pulse profile is quite complex with a double-peaked component followed by weaker third and fourth component. The latter peaks could be interpreted as interpulses[23], however, these peaks are not separated by 180° in phase, therefore making any interpretation difficult at this stage. MSPs unlike normal pulsars tend to have more complicated pulse profiles on average displaying more pulse components (see for e.g. Yan et al. 2011). Some examples of MSPs in this category include 47TucQ (Camilo et al., 2000) and PSR J1536−4948 (Bhattacharyya et al., 2021). A normal pulsar which also shows quite a complicated pulse profile is PSR B1237+25 (see e.g. Srostlik & Rankin 2005). Multiple theoretical models have been proposed in order to explain complex pulse profiles (for e.g. Rankin 1983; Lyne & Manchester 1988; Kramer et al. 1994; Gupta & Gangadhara 2003). However, no single model can yet explain all the observed diverse range of pulse profiles. Studies of the integrated pulse-profile are useful because the size and shape of the pulse components highlight the geometry of emission regions in the pulsar magnetosphere. However, in order to investigate these properties further for PSR J1753−24 we need polarimetric data. This has not been performed yet and is expected to be done in the future once we have sufficient TOAs for a phase-connected timing solution. Additionally, this pulsar has also been recently confirmed to be in a binary orbit. Using detections from followup search-mode observations from Parkes, we used the program FITORBIT to solve the orbit of this pulsar to an orbital period of 70.7 days with a projected semi-major axis value of 5.6(1) lt-s. This gives us the minimum companion mass of this pulsar which comes out to $\sim 0.043\,\mathrm{M}_\odot$. Our current best ephemeris of this pulsar can be found in table 3.8. This ephemeris was derived from a total of 9 detections spanning across two months. The orbit of this pulsar appears to be nearly circular. Currently, we do not have enough detections to place limits on its orbital eccentricity and given our large position uncertainty of the order $14'$, we have currently set the spin-frequency derivative $\dot{\nu}$ in our ephemeris to be equal to zero. A full phase-connected timing solution of this pulsar along with calibrated flux density measurements and polarisation observation results will be the subject of a future publication.

---

[23] Interpulses are secondary pulses apart from the main pulse usually separated by 180° in phase.



Search Information

Candidate: 11.76ms_Cand
Telescope: Parkes
$\mathrm{Epoch_{topo}}$ = 56259.23045138889
$\mathrm{Epoch_{bary}}$ = N/A
T_{sample} = 6.4e−05
Data Folded = 67581600
Data Avg = 1388
Data StdDev = 27.55
Profile Bins = 64
Profile Avg = 1.466e+09
Profile StdDev = 2.831e+04

$\mathrm{RA_{J2000}}$ = 17:43:43.5000 $\mathrm{DEC_{J2000}}$ = −24:27:25.9000

Best Fit Parameters

$\mathrm{DOF_{eff}}$ = 56.01 χ^2_{red} = 6.343 P(Noise) < 1.98e−50 (14.9σ)
Dispersion Measure (DM; pc/cm^3) = 196.695
P_{topo} (ms) = 11.75786550(69) P_{bary} (ms) = N/A
P'_{topo} (s/s) = 1.8(1.2)x10^{-12} P'_{bary} (s/s) = N/A
P''_{topo} (s/s^2) = 0.0(1.9)x10^{-15} P''_{bary} (s/s^2) = N/A

Binary Parameters

P_{orb} (s) = 25483.374941 e = 0.000000
$a_1\sin(i)/c$ (s) = 0.000000 ω (deg) = 0.000000
T_{peri} = 56258.97983020520

Figure 3.10: Discovery plot of PSR J1743−24 found in the LOWLAT beam 2012-11-28-05:31:51/11. The orbital template that was used to detect the pulsar can be found in the top right under ‘Binary Parameters’.

Table 3.8: Best fitting orbital parameters of PSR J1743−24.

Pulsar	J1743−24
Right Ascension, α (J2000)	17:43:43.50
Declination, δ (J2000)	−24:27:25.90
Spin Frequency, ν (s^{-1})	85.05187(2)
1st Spin Frequency derivative, $\dot{\nu}$ (Hz s^{-2})	0(0)
Reference Epoch (MJD)	56062(5)
Dispersion Measure, DM (pc cm^{-3})	197.359
Binary Parameters	
Binary Model	BT
Projected Semi-major Axis, $x_{\rm p}$ (lt-s)	5.6(1)
Longitude of Periastron, ω (deg)	248.929(5)
Epoch of passage at Periastron, T_0 (MJD) ...	56062.043(7)
Epoch of passage at Ascending Node, $T_{\rm asc}$ (MJD)	56083.860(5)
Orbital Period, P_b (days)	70.7(1)
Eccentricity, e	$4.7(1)\times10^{-7}$
Derived Parameters	
Spin Period, P (s)	$1.1757531(2)\times10^{-2}$
1st Spin Period derivative, $\dot{P}$ (s s^{-1})	−0(0)
Mass Function, $f(M_{\rm p})$ ($M_\odot$)	$3.87(1)\times10^{-5}$
Minimum companion mass, $M_{\rm c,min}$[a]($M_\odot$)	$4.32(1)\times10^{-2}$
Median companion mass, $M_{\rm c,med}$[b]($M_\odot$)	$5.00(4)\times10^{-2}$

[a] $M_{\rm c,min}$ is calculated assuming the orbital inclination angle $i = 90$ and pulsar mass M_p is 1.4 $M_\odot$.

[b] $M_{\rm c,med}$ is calculated assuming the orbital inclination angle $i = 60$ and pulsar mass M_p is 1.4 $M_\odot$.

3.3.3.1 Discussion of PSR J1753−2819 and PSR J1743−24

Before, we discuss the properties of PSR J1753−2819 and PSR J1743−24, we start with the standard spin, orbital and mass distribution of binary MSPs which is useful in order to understand the uniqueness of both these pulsars. MSPs in binary orbits are known to have a wide range of companion stars. Vast majority of these binary pulsars have extremely circular orbits with a white-dwarf (WD) or low-mass companion ($m_c \leq 0.06 M_\odot$) and they typically tend to be highly recycled ($P \leq$ 10ms). Depending on the mass of the companion, these binaries can be categories into two different populations. One is the so called very low-mass binary pulsars (VLMBPs; Freire et al. 2003) which have spin-periods in the order of ∼ 4ms, companion masses $m_c \leq 0.03 M_\odot$, orbital period $P_{\rm orb} \leq$ 10hrs. Another category is the so called low-mass binary pulsars (LMBPs; Phinney & Kulkarni 1994; Edwards & Bailes 2001). They typically tend to have companion masses in the range of $0.1 - 0.5 {\rm M}_\odot$, orbital period of few days and rapid spin-periods in the order of few milliseconds. There are some notable exceptions with longer spin-periods like PSR J1822-0848 (Lorimer et al., 2006), PSR J1932+1500 (Lyne et al., 2017) among others. Most MSPs in a binary orbit tend to fall under the LMBPs category and they typically have He-WD companions. However MS stars and CO WD stars are also possible. Intermediate-spin period pulsars (10ms $\leq P \leq$ 20 ms) tend to have heavier CO WD companions and pulsars with spin period greater than 20ms could have either a neutron-star (NS) or a WD or a main-sequence star (MS; Johnston et al. 1992; Stairs et al. 2001) companion. These are typically called Intermediate-mass binary pulsars (IMBPs). We would like to stress here that this is a highly simplified view with several overlapping exceptions and the boundaries between these three categories are not fixed.

Our timing solution confirms that PSR J1753−2819 is a binary pulsar with an orbital period of $\sim$ 9.3 hrs. We obtained a projected semi-major axis of 0.20954(5) lt-s. The mass function of the pulsar $f = 6.585 \times 10^{-5} \mathrm{M}_{\odot}$. Using this, and assuming a pulsar mass of $1.4\,\mathrm{M}_{\odot}$, an inclination angle $i = 90°$, we can calculate the minimum companion mass of this pulsar which comes out to be $\sim 0.052\,\mathrm{M}_{\odot}$. Assuming orbital inclinations are randomly distributed in nature, by setting inclination angle $i = 60°$, we can get a median companion mass limit which comes out to be $\sim 0.06\,\mathrm{M}_{\odot}$. These values indicate that its likely that the pulsar is in orbit around a low-mass companion star. Similar arguments can also be made for PSR J1743−24 which has a similar minimum and median companion mass of $\sim 0.043\,\mathrm{M}_{\odot}$ and $\sim 0.05\,\mathrm{M}_{\odot}$ respectively but is part of a much longer orbit of 70.7 days. Though unlikely, it is also possible that the companion of both PSR J1753−2819 and PSR J1743−24 could be a low-mass Helium White-Dwarf (He-WD) star. In order to obtain a WD mass of at least $0.154\mathrm{M}_{\odot}$[24], assuming the same pulsar mass as above, this would require an inclination angle lesser than $\sim 20.5°$ for PSR J1753−2819 and $\sim 17.0°$ for PSR J1743−24 in order to obtain our projected semi-major axis measurements. These have the probability of $\sim$ 0.063 and $\sim$ 0.043 respectively when drawn from a uniform distribution of inclination angles. Assuming that the companion is indeed the more likely low-mass companion star, one of the interesting aspects of PSR J1753−2819 and PSR J1743−24 is its intermediate spin-period value of 18.62 ms and 11.76 ms respectively. Using, the standard picture of MSP binary distribution described above, we would expect both these pulsars to be much more recycled. The rarity of such pulsars can be demonstrated in two steps. First, we plot the orbital period versus the minimum companion mass of known MSP binaries[25] along with both these pulsars, this is shown in the top part of figure 3.11. We can clearly see at least three distinct populations. Colors and markers in this diagram indicate the nature of the companion star. In the left side of the diagram

---

[24] This is the lowest WD mass we obtained in the sample provided by the Extremely Low-Mass White Dwarf Survey (ELM, Brown et al. 2020)-a systematic spectroscopic survey of nearby ($\leq$ 2 kpc) low-mass He-White Dwarfs

[25] We have only included Galactic field MSP binaries for our discussion, i.e these pulsars are not associated with a Globular Cluster. Pulsars in GCs tend to have a more complicated evolutionary history due to exchange interactions.

($m_{\mathrm{c,min}} \leq 0.03 M_{\odot}$), we have the VLMBPS, which are highly recycled pulsars with orbits lesser than a day and they have ultra-light companion stars. This is followed by LMBPS which composes mostly of He-WD companions. There is a weak trend between orbital period and minimum companion mass for PSR-He-WD binaries. Finally, on the right hand side, we have the IMBPS which have a combination of CO WD companions, DNS and MS companions. Both these pulsars occupy the region between VLMBPS and LMBPs which gives us important clues regarding the true boundaries between the two classes. A final classification of these pulsars into one of these categories can only be done once we identify the nature of the companion star or calculate the mass of the companion which would require the measurement of two Post-Keplerian parameters. In the bottom part of figure 3.11, we plot the minimum companion mass versus the spin-period of the pulsar where the uniqueness of both these pulsars becomes more evident. We can see the PSR J1753−2819 and PSR J1743−24 occupy the bottom right region of the diagram where only four other binaries are known. Some comparable examples include PSR J1502−6752 (Keith et al., 2012) which has a spin-period of 26.7 ms and a minimum companion mass of $0.02 M_{\odot}$. PSR J1727−2951 (Cameron et al., 2020) is a candidate black-widow binary system which unlike a typical black-widow pulsar is only partially recycled and shows no evidence of eclipses. PSR J1744−3922 is a 172-ms pulsar which is in a 4.6 hour orbit and a companion mass of $0.08\mathrm{M}_{\odot}$ and is known to exhibit highly variable pulsed radio emission (Breton et al., 2007). An extreme example is B1831−00 which is a pulsar with a spin-period of 520.95-ms in a 1.8 day orbit with a minimum companion mass of $0.063\mathrm{M}_{\odot}$ (Hobbs et al., 2004). The spin, orbital periods and companion mass ranges of these six pulsars are likely inconsistent with the standard evolutionary models and could likely be part of a separate population of pulsars with intermediate companion masses and crucially, slower (and in some cases much slower) spin periods.

3.3.4 Other noteworthy discoveries

PSR J1609−5225 is a normal pulsar rotating once every 231.8080(7) ms and has the highest detected DM in our discovery sample of 993(10) $\mathrm{pc\,cm^{-3}}$. Timing observations of this pulsar have currently gained priority due to the redetection of this pulsar in the ongoing MGPS-L survey using MeerKAT which gave us an improved position measurement of RA(J2000): 16:09:02.41, DEC(J2000): -52:25:18.6, Galactic longitude Gl = 330.516° and Galactic latitude Gb = -0.468°. Based on the NE2001 and YMW16 electron density models, this pulsar is located at a distance of 14.32 kpc and 7.20 kpc respectively. Since we do not yet have independent distance measurements, we cannot yet confirm which of these distance estimates are more reliable. If the NE2001 DM distance does prove to be more reliable, then it places this pulsar at a distance greater than our distance towards the Galactic centre. The integrated pulse profile of this pulsar shows a strong scattering tail. As discussed earlier in section 1.5.2, if the pulsar signals pass through dense regions of plasma (for e.g. HII regions) on their way to our receivers, then these density fluctuations act as a scattering screen and scatter the profile due to multipath propagation of light which causes the signals to

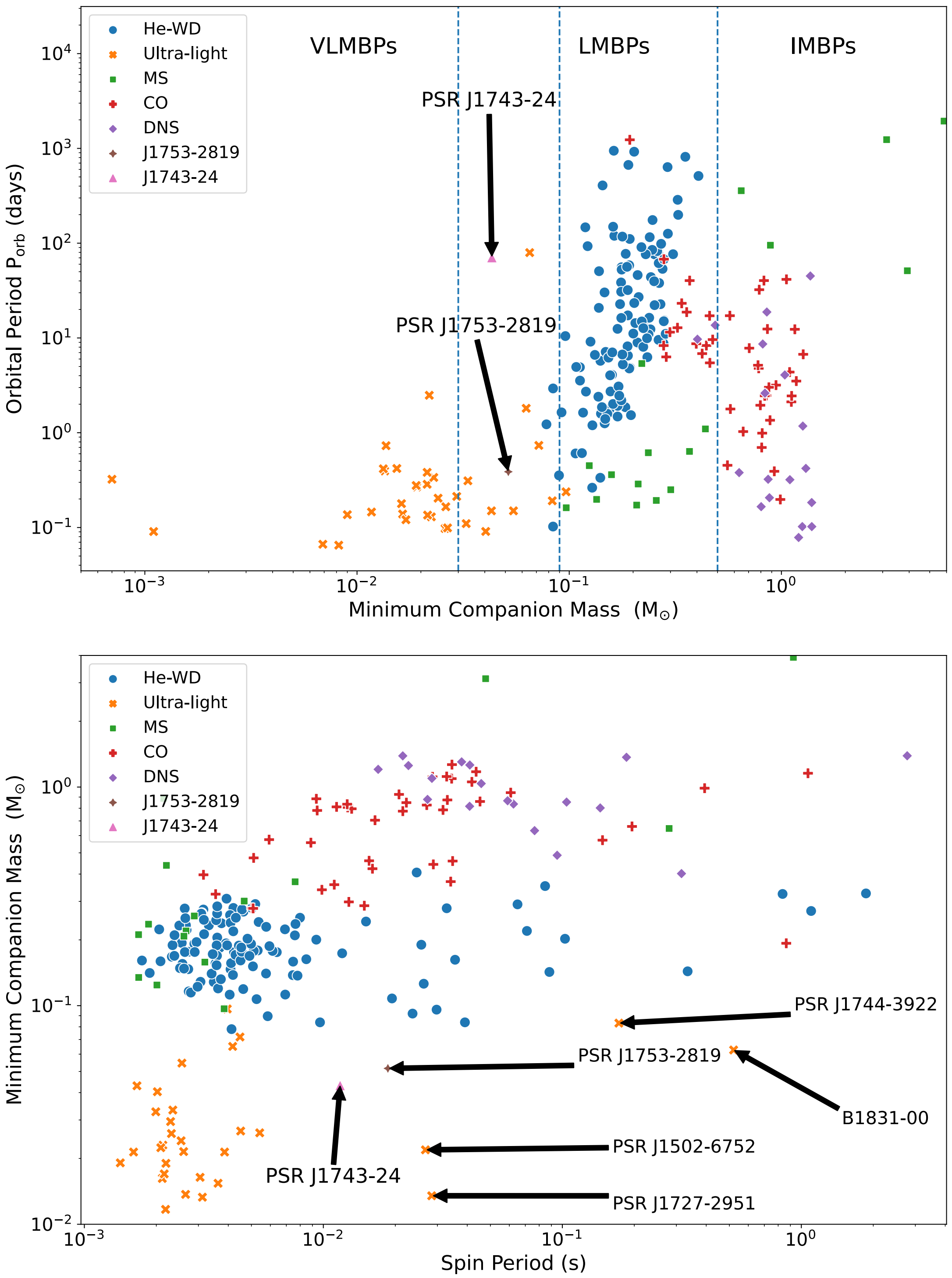


Figure 3.11: Top- Orbital Period versus minimum companion mass for all known Galactic-field MSP binaries along with PSR J1753−2819 and PSR J1743−24. The markers and colors highlight the nature of the companion star. We have also divided the diagram into regions occupied by VLMBPS, LMBPS and IMBPS respectively. Bottom - Spin Period versus minimum companion mass for all known Galactic-field MSP binaries along with PSR J1753−2819 and PSR J1743−24. Points labelled by the pulsar name in the bottom right side are the exceptions to the standard distribution of MSP binaries. See text for further discussion.

arrive at different times at our receivers. This broadens the pulse shape creating a long tailed exponentially decaying trailing edge component of the pulse profile. We modelled for this by fitting an exponential function to the trailing edge of the pulse profile and our best model gives us a scattering timescale $\tau = 77 \pm 25\,\mathrm{ms}$. These results are shown in figure 3.12. Our model estimates were made on a search-mode observation of PSR J1609−5225. Therefore, the uncertainty in our scattering model is relatively high. This can be resolved once we phase-connect this pulsar and stack pulse profiles from different observations together in order to get a high S/N pulse profile template. This is expected to be done in the near future once we commence the timing campaign of this pulsar. Our results are below the predictions for scattering timescale for these coordinates and DM derived from both the NE2001 and YMW16 electron density models $\tau_{sc,\mathrm{NE2001}} = 149.7 \pm 3.2\mathrm{ms}$ and $\tau_{sc,\mathrm{YMW16}} = 307.6 \pm 13.2\mathrm{ms}$. These model predictions are only expected to be rough estimates as these values depend a lot on the light of sight and should mostly be used as upper limits. Future observations of this pulsar will be aided by observing with a wide frequency bandwidth using the PARKES UWL receiver which will help to calculate the scattering spectral index α. Pulsars like PSR J1609−5225 are useful to improve our understanding of the Galactic electron density distribution and would be a useful addition to the sample of pulsars that are more relevant for ISM related studies.

PSR J1142-62 is a normal pulsar with a detected barycentric spin period in the discovery observation of $P = 1573.494(14)\,\mathrm{ms}$. This pulsar is located at Galactic longitude Gl = 295.05° and Galactic latitude Gb = -0.55°. The distance to this pulsar based on the NE2001 and YMW16 electron density models are 5.21 kpc and 2.49 kpc respectively. During our follow-up confirmation observation, we detected an apparent spin-up in the barycentric spin-period $P = 1573.490(14)$ ms indicating that this could be yet another binary candidate. At the current stage, due to the uncertainties in our spin-period measurements, we cannot yet disentangle if PSR J1142-62 is an isolated pulsar or a pulsar in a binary orbit. If this pulsar does happen to be in a binary, then it will be part of a minority population as only 2 % of normal pulsars reported in PSRCAT[26] are known to be in a binary orbit. We plan to regularly monitor this pulsar using Parkes in the near-future. Further follow-up observations will be crucial in order to reveal the true nature of this pulsar.

3.3.5 ClassA Candidates

Besides the confirmed pulsars mentioned above, three class A candidates (J1546−54, J1630−44 and J1723−38) await further confirmation. Our candidate detection plots are shown in figures 3.13, 3.14 and 3.15 respectively. A list of the barycentric spin period, dispersion measure, sky coordinates, the HTRU pointing, beam parameters and DM dependent distance estimates can be found in table 3.9. J1546−54 is a 1466.829(6) ms candidate, detected with a DM = 307(14) $\mathrm{pc\,cm^{-3}}$. This candidate was detected at zero acceleration and is likely an isolated pulsar. Our current position measurements based on the location of the LOWLAT beam places it at Galactic longitude Gl = 326.6° and

---

[26] https://www.atnf.csiro.au/research/pulsar/psrcat/

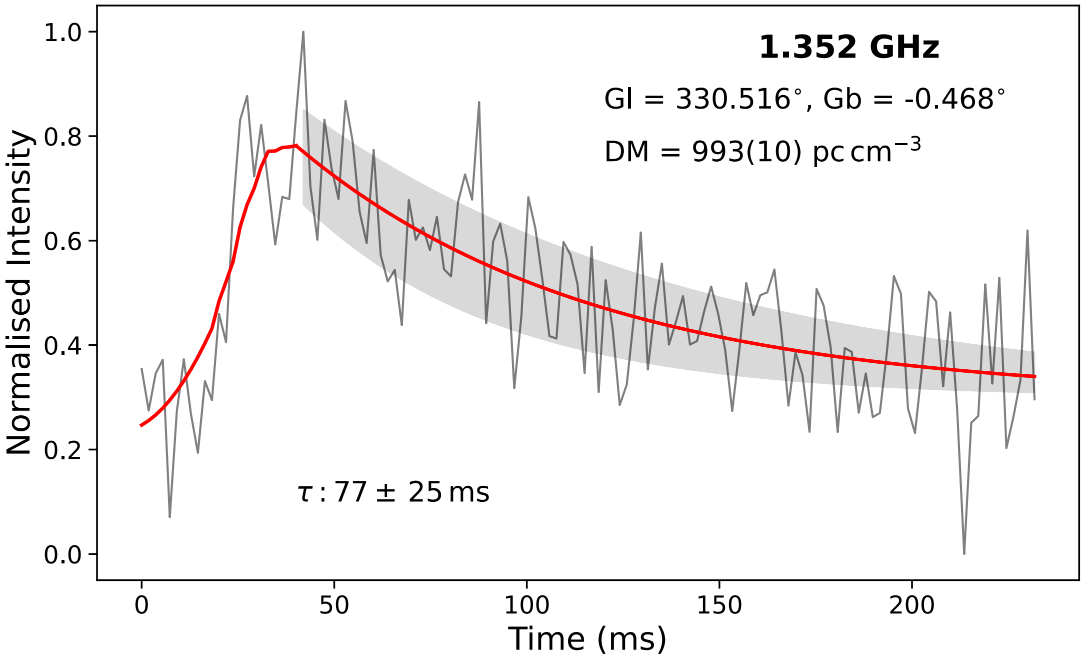


Figure 3.12: Integrated pulse profile of PSR J1609−5225 at a central frequency of 1.352 GHz plotted in light grey. Red curve shows the best fit scattering model. Dark gray intervals around the red curve display the uncertainty in our best-fit parameters. We report here the scattering time scale τ which was calculated from a LOWLAT search-mode observation with an effective bandwidth of 340 MHz.

Galactic latitude Gb = 0.1°. The remaining two candidates (J1630−44 and J1723−38) were detected with a non-zero acceleration and are likely candidates in a binary orbit. Using the orbital parameters of our template that made the detection, we get a maximum line-of sight acceleration value of $140.2\,\mathrm{m\,s^{-2}}$ and $37.1\,\mathrm{m\,s^{-2}}$ respectively. Our current position measurement for the candidate J1630−44 is Gl = 338.7°, Gb = 2.7° and for J1723-38 is Gl = 349.1°, Gb = -1.6°. Both these are normal pulsar candidates with a spin-period of 373.315(2) ms and 237.205(1) ms respectively. J1630−44 and J1723−38 were exclusively found in the template-bank pipeline and we are fairly certain that they are real due to their measured detection significance being above 9 sigma or S/N > 10. Additionally, the DM dependent distance to these candidates based on the YMW16 and NE2001 electron density models are below the maximum predicted distance assuming that these sources are within our Galaxy, therefore increasing our confidence in these candidates. However, follow-up observations of all these candidates have so far yielded no detections. This is most likely caused due to the effect of RFI in broadband observations of the UWL with the MEDUSA backend. Several of our confirmed pulsars for e.g. PSR J1142−62, PSR J1737−33 among others were missed in the FFT in followup observations with the UWL. We were able to confirm them by directly folding the pulsar with the known spin-period, manually cleaning the folded archive and doing a small period offset search using PDMP. While this is an effective technique, it works only for isolated pulsars. Confirming slow pulsars in a binary orbit with UWL remains a challenge, and will likely need new techniques to mitigate the effect of RFI in search-mode observations. The alternative less likely scenario is that J1630−44 and J1723−38 are false-alarm candidates caused due to noise lining up by chance. However, given the high detection significance of these candidates[27] we think this scenario is unlikely. Yet another possibility for our lack of detections could be emission instability if these candidates turn out to be intermittent pulsars (discussed earlier in section 1.6.4). These questions can only be answered with more follow-up observations which is one of the immediate near-term goals of this project.

---

[27] See the third line in the top right of the detection plots for the probability that these could be noise assuming a single trial search. Probability of noise P(Noise) < 2e-20 even in the most conservative scenario.

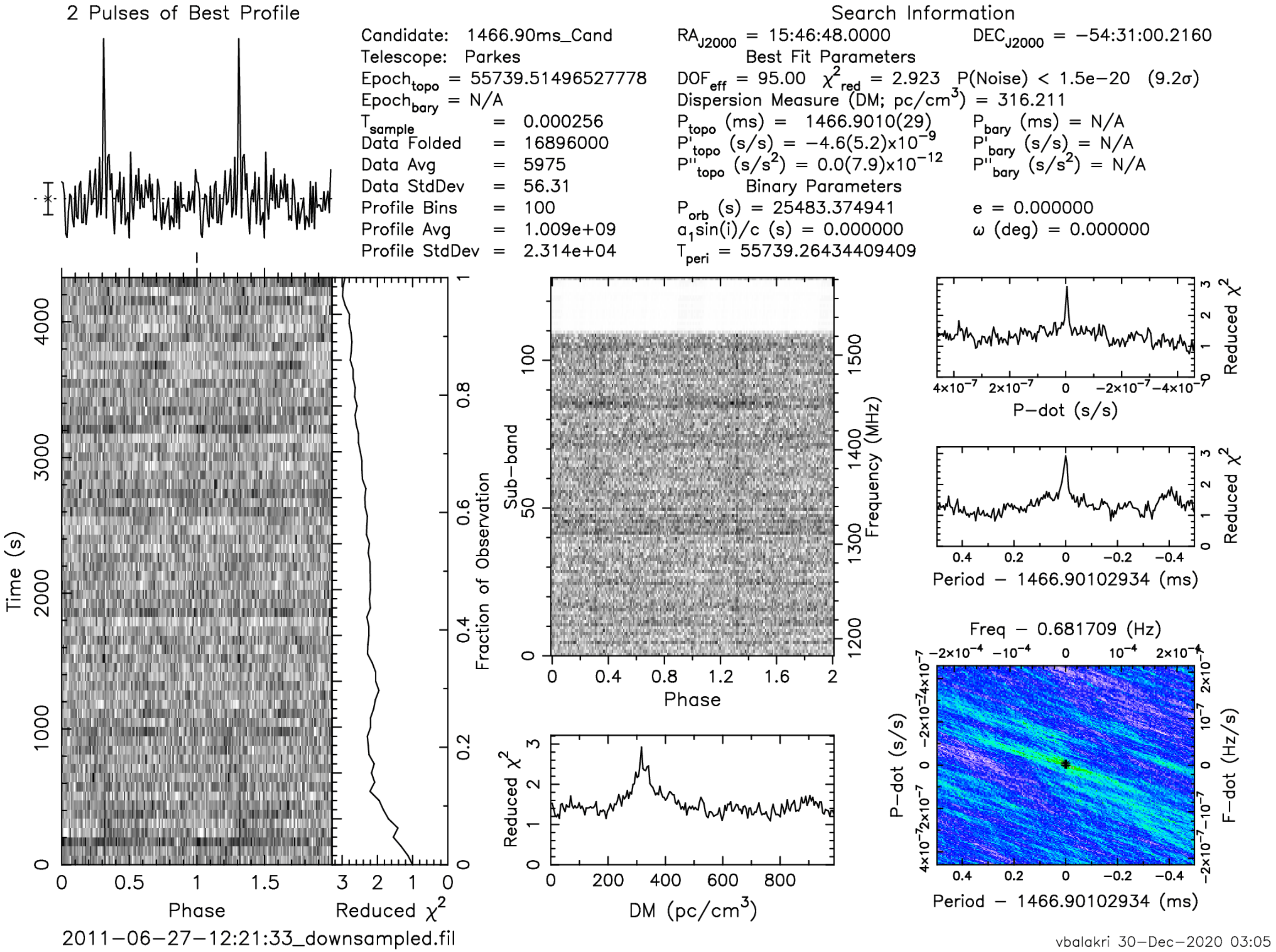
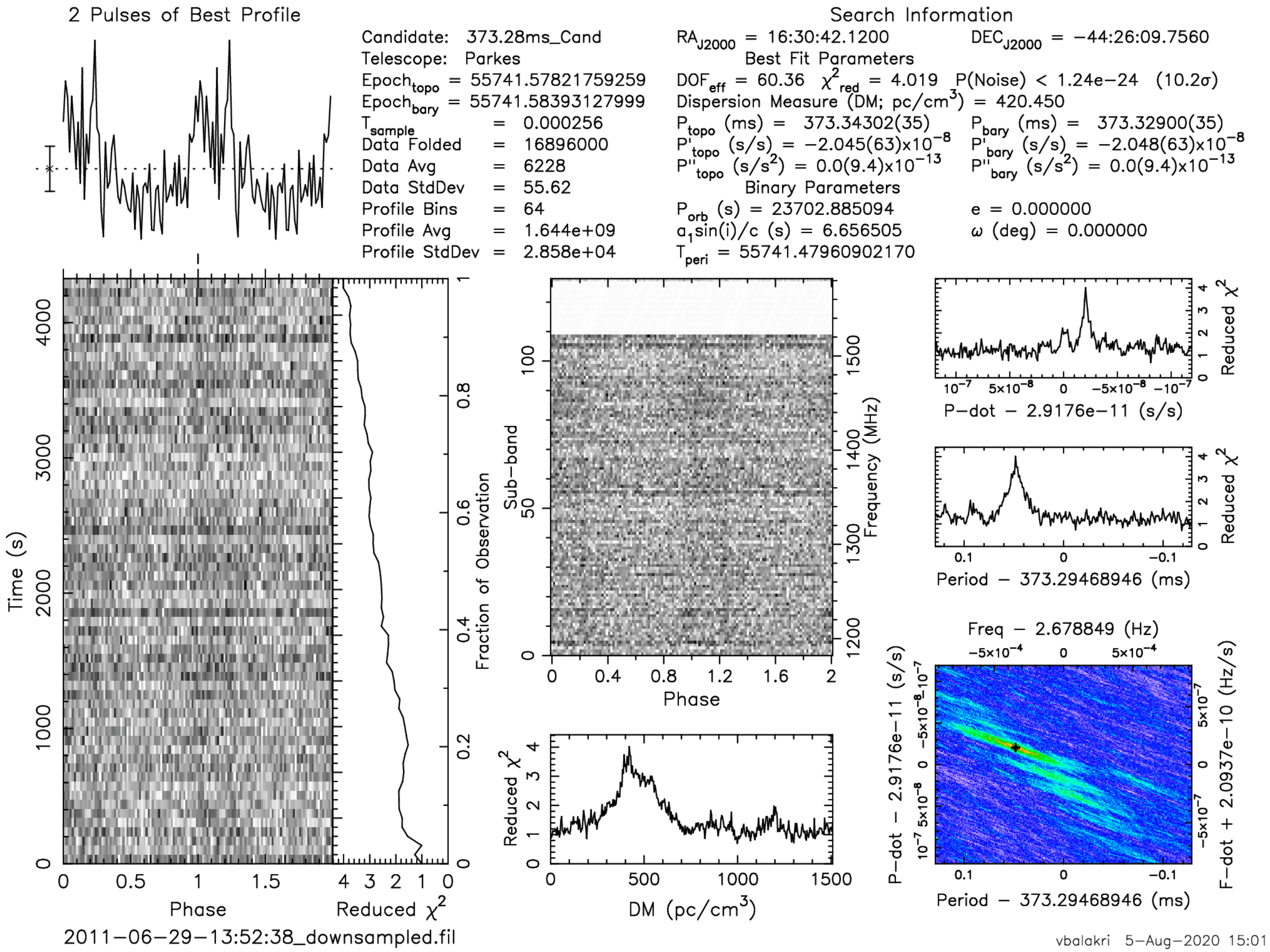


Figure 3.13: Detection plot of one of our Class A pulsar candidate J1546−54 found in the LOWLAT beam 2011-06-27-12:21:33/13. This signal was detected with zero acceleration and its likely that J1546−54 is an isolated pulsar. This candidate has been followed up six times using Parkes between MJD 59212 and 59399 with no detections.



```
Candidate:  373.28ms_Cand            RA_J2000 = 16:30:42.1200              DEC_J2000 = −44:26:09.7560
Telescope:  Parkes                            Best Fit Parameters
Epoch_topo = 55741.57821759259      DOF_eff = 60.36  χ²_red = 4.019  P(Noise) < 1.24e−24  (10.2σ)
Epoch_bary = 55741.58393127999      Dispersion Measure (DM; pc/cm³) = 420.450
T_sample       =  0.000256          P_topo (ms) =  373.34302(35)     P_bary (ms) =  373.32900(35)
Data Folded    =  16896000          P'_topo (s/s) = −2.045(63)x10⁻⁸   P'_bary (s/s) = −2.048(63)x10⁻⁸
Data Avg       =  6228              P''_topo (s/s²) = 0.0(9.4)x10⁻¹³  P''_bary (s/s²) = 0.0(9.4)x10⁻¹³
Data StdDev    =  55.62                      Binary Parameters
Profile Bins   =  64                P_orb (s) = 23702.885094          e = 0.000000
Profile Avg    =  1.644e+09         a₁sin(i)/c (s) = 6.656505         ω (deg) = 0.000000
Profile StdDev =  2.858e+04         T_peri = 55741.47960902170
```

Figure 3.14: Detection plot of the pulsar candidate J1630−44 found in the LOWLAT beam 2011-06-29-13:52:38/01. The orbital template that was used to detect the pulsar can be found in the top right under 'Binary Parameters'. Based on these orbital values, we calculate the maximum line of sight acceleration of this signal to be 140.2(2) ms^{-2}. This candidate has been followed up 9 times at Parkes between MJD 59063 and 59399 with no detections yet.

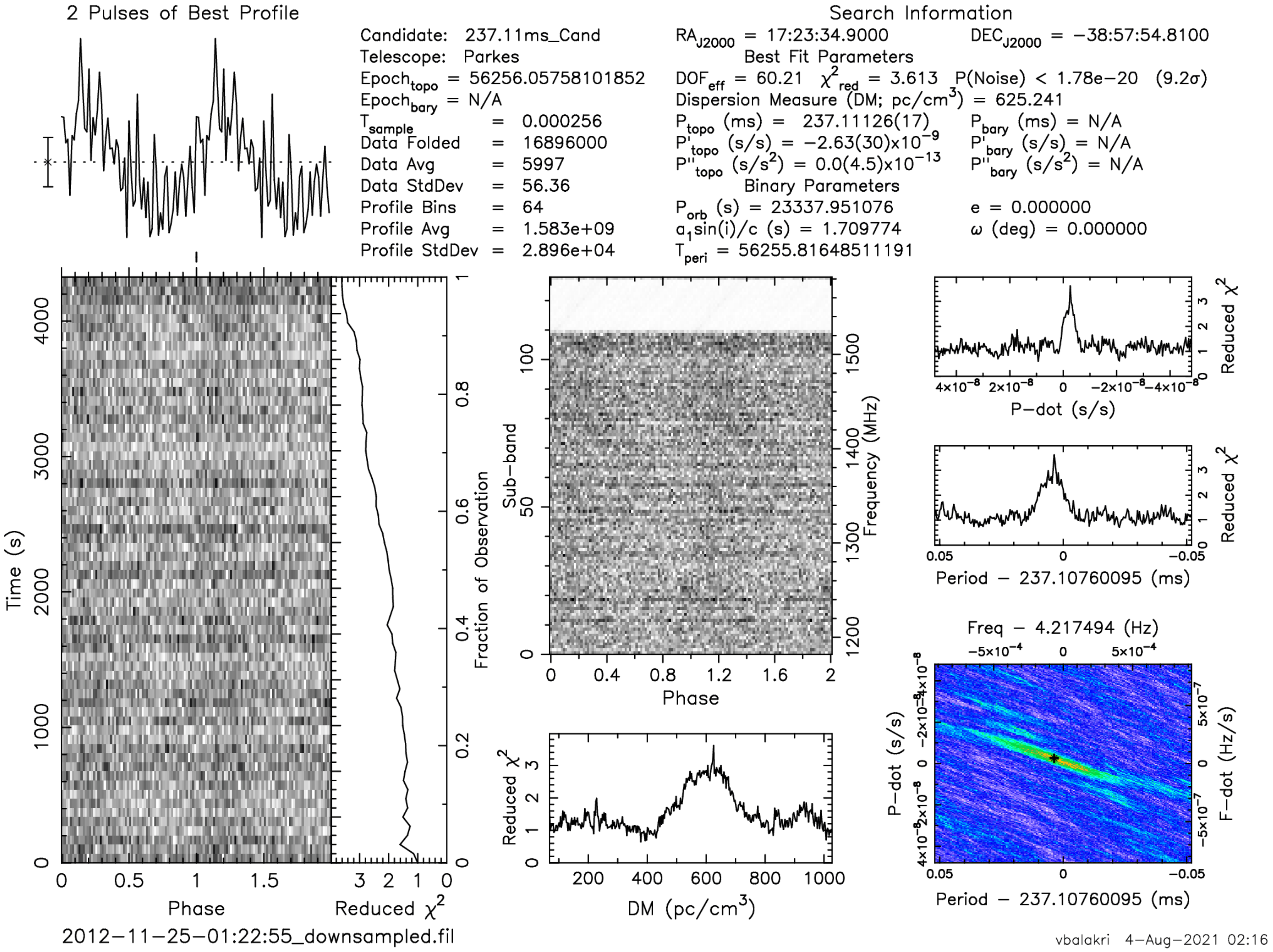


Figure 3.15: Detection plot of the pulsar candidate J1723−38 found in the LOWLAT beam 2012-11-25-01:22:55/12. Based on the orbital template that made the detection we calculate the maximum line of sight acceleration of this signal to be 37.1(5) ms^{-2}.

Table 3.9: A list of Period, Dispersion measure and sky coordinates of our Class-A candidates defined as candidates that are highly likely to be real due to the high S/N of our detection. Despite multiple followup observations, none of these have been confirmed yet. We also report the DM dependent distance estimate of these pulsar candidates based on the NE2001 and YMW16 electron density models. See main text for further discussion.

PSRJ	pointing/beam	RA (J2000) (hms)	Dec (J2000) (dms)	gl (°)	gb (°)	P (ms)	DM ($\mathrm{pc\,cm^{-3}}$)	S/N	PEPOCH MJD	D_{YMW16} (kpc)	D_{NE2001} (kpc)
J1546−54	2011-06-27-12:21:33/13	15:46:48.00	-54:31:00.2	326.6	0.1	1466.829(6)	307(14)	11.16	55739.545224	4.44	4.73
J1630−44	2011-06-29-13:52:38/01	16:30:42.12	-44:26:09.75	338.7	2.7	373.315(2)	421(21)	13.20	55741.604721	18.65	7.58
J1723−38	2012-11-25-01:22:55/12	17:23:34.90	-38:57:54.81	349.1	-1.6	237.205(1)	585(15)	12.43	56256.078213	4.05	3.66

3.3.6 Redetection of an additional 50 new pulsars and comparison to the original discovery population of LOWLAT

A total of **70** new confirmed pulsars have been found so far using the template-bank pipeline. In section 3.3.1 we reported the discovery parameters of **20** pulsars. In this section we report the remaining **50** pulsars which were independently detected by our pipeline but were initially found by other ongoing reprocessing searches in LOWLAT. All these **70** pulsars were missed by the first-pass search pipeline (Ng et al., 2015; Cameron et al., 2020). 47 of these pulsars were initially found by a GPU-Accelerated Acceleration-Search pipeline (Sengar et al., prep). These searches are currently been carried out on the full-length 72 minute observation of LOWLAT at the native time resolution ($t_{\text{samp}} = 64\,\mu\text{s}$) as well as observations decimated by a factor of 4 ($t_{\text{samp}} = 256\,\mu\text{s}$). A further 3 pulsars were found by an ongoing reprocessing using a modified version of FFA Pipeline RIPTIDE[28] (Morello et al., 2020; Wongphechauxsorn et al., prep). We give a list of the barycentric spin period, dispersion measure, sky coordinates and the corresponding HTRU pointing and beam where the pulsar was found using our pipeline in table 3.10. We also present the DM dependent distance based on the 'NE2001' model (Cordes & Lazio, 2002) and the YMW16 model (Yao et al., 2017). Further analysis of these pulsars including timing solutions will be the subject of a future publication (Sengar et al., prep; Wongphechauxsorn et al., prep). In order to identify why such a large number of pulsars were discovered in the reprocessing which will aid future search strategies for LOWLAT and other pulsar surveys, we present here a comparison of the S/N, DM and Galactic longitude distribution of the 70 new pulsars presented in this work (hereafter referred to as 'Current Discoveries') to the 100 previously discovered pulsars (hereafter referred to as 'Previous Discoveries') in LOWLAT reported in Ng et al. (2015); Cameron et al. (2020). The first-pass pipeline was used to do an acceleration search using a segmented search approach (search range given earlier in table 3.1), in order to be sensitive to relativistic binary pulsars. It had 15 independent search components, candidates from which were grouped together, and ranked by FFT spectral S/N with a cutoff at S/N equals 8.0. Approximately the first 1000 ranked candidates per beam were then folded and inspected by eye. While, this search was overall successful given that it found one of the most relativistic binary pulsars known till date - PSR J1757−1854 (Cameron et al., 2018), this approach has its drawbacks in terms of missing several low spectral S/N pulsars. Since all the reprocessing searches, were entirely focused on the full-length 72-minute observation, and with gains in computational power, we were able to have a lower spectral S/N threshold of 7 and fold a factor of 2-3 more candidates per beam which enabled us to find pulsars which were ranked well below the cut-off threshold of the initial first-pass pipeline. This increase in number of pulsar discoveries at low S/N is demonstrated in figure 3.16 where we show a histogram and cumulative distribution function (CDF) of the spectral and folded S/N of the current pulsar discoveries to the ones previously reported in Ng et al. (2015); Cameron et al. (2020). 48 out of the 70 new pulsars had a spectral S/N greater than or equal to 8 and in principle should have been detected by

---

[28] `https://github.com/v-morello/riptide`

past searches. The remaining 22 pulsars were below the FFT S/N threshold of 8.0 and its unlikely that these candidates survived the FFT thresholding stage of the previous pipeline. There are primarily four reasons to explain why the first 48 pulsars were missed. The first-pass sigproc-based pipeline before commencing its search decimated the observations by a factor of 4 ($t_{\mathrm{samp}} = 256\,\mu\mathrm{s}$) for purely computational reasons using the software SIGPYPROC[29]. However, during this process the data were incorrect re-scaled back to 2-bits which effectively reduced the bit-depth of the data and thereby reducing the sensitivity of our observations. This has since been fixed in software for the current reprocessing by outputting a 4-bit observation after downsampling. The second reason was the DM tolerance used by the sigproc dedispersing program DEDISPERSE_ALL has a default DM tolerance of 1.25 compared to 1.11 in PEASOUP. This effectively leads to more DM trials between two values for the reprocessing searches and thereby reducing the S/N lost due to sampling at an incorrect DM. The third reason was the limited amount of candidates that could be folded due to computation and the implementation of the segmentation search pipeline which had 15 independent search components (Ng et al., 2015). We were able to find these new pulsars due to the increase in computational power available now, the development of new GPU-based pipelines and fixing the downsampling software bug in SIGPYPROC. Finally, the fourth reason is human error. During the processing of a large scale pulsar survey like LOWLAT, it is expected that a few pulsars will be missed due to human error during inspection. This is likely to have been the case with the PSR J1814−18 which had a spectral and folded S/N greater than 30. This may have also played a role in some of other pulsar discoveries reported here. Similar to spectral S/N, we see a trend of increased discovery at lower folded S/N as well with the median folded S/N of the current discovery population at 11.5 compared to a median S/N of 14.1 for the previously discovered population in LOWLAT. Using a two-sample Kolmogorov–Smirnov (KS) test, we get p-values of 5×10^{-8} and 9×10^{-8} for the spectral and folded S/N distributions respectively and therefore, the null hypothesis (i.e the distribution of S/N from the current and previous discoveries are identical) can be ruled out. Given the same survey parameters, and assuming that the period and pulse width distribution of both population of pulsars is similar, a lower S/N could either mean that we are finding low luminosity pulsars closer to us or comparable luminosity pulsars located further away from us. For this comparison, ideally we would like to have independent distance measurements to each pulsar, and calibrated flux density measurements. However, in the absence of these two measurements, we can use the detected DM of each pulsar as a proxy for distance and compare the distributions for both populations. This is shown in figure 3.17. We see a clear statistically significant trend in the discovery of more pulsars at higher DM. The median detected DM of the previously discovered population of pulsars is 326 compared to 470 in the current discovery population. Using a KS test, we get a p-value of 0.0016, and therefore the null hypothesis that the DM distributions of both populations are identical can be safely ruled out. Based on these results, it is likely that we are probing a phase-space of pulsars located further away from us. Additionally, in order to compare

---

[29] https://github.com/ewanbarr/sigpyproc

the spatial distribution of both populations of pulsar discoveries, in figure 3.18 we show a histogram of their Galactic longitude (gl) distribution. The scatter plot overplotted on the histogram shows the detected DM of each pulsar. The distributions are broadly consistent with each other. There is a larger than average number of new discoveries between galactic longitude 8 and 13 degrees, however this trend is not statistically significant. It is likely that this region by coincidence happened to have more low S/N pulsars that were undetected by past searches. Using the KS test, we get a p value of 0.49 and therefore the null hypothesis that the gl distributions of both populations are identical cannot be ruled out. Besides, the 70 pulsars given here, an additional 22 pulsars is expected to be announced in Sengar et al. (prep) in the near future. These were not detected using our pipeline. There are primarily two reasons for this. First, is that all the remaining undetected pulsars were discovered by a mild acceleration search on the native time resolution of LOWLAT ($t_{\mathrm{samp}} = 64\,\mu\mathrm{s}$) whereas our search was done on decimated data at 256 µs due to the larger number of binary trails in our search. These pulsars were missed due to the FFT detection significance at 256 µs being below our cutoff limit of 7.0. A second reason is our high threshold of $P_{\mathrm{spin}} = 13\,\mathrm{ms}$ in the template-bank pipeline. Therefore, we are only partially sensitive towards isolated pulsars or binaries below this limit. Besides, the new discoveries presented so far, we also report an additional 60 unique known pulsars that were missed by the first pass pipeline in table A.1. This list was compiled based on the redetection list provided by Ng et al. (2015); Cameron et al. (2020). We think these pulsars were likely missed due to the reasons mentioned above and human error while inspecting pulsar candidates. Our results show that for future pulsar surveys using the SKA, it will be crucial to fold a larger number of pulsar candidates per beam in order to maximise the number of discoveries from a survey. However, this comes at the cost of higher computation. Alternatively, more work needs to go into better RFI-mitigation algorithms or clever heuristics which can reject likely false-positive candidates from the FFT. A second takeaway from these results is that a higher time resolution is crucial to find low S/N pulsars. The fact that a total of 22 pulsars can be exclusively found in the native time resolution of $t_{\mathrm{samp}} = 64\,\mu\mathrm{s}$ is quite significant, therefore any future pulsar search strategies that rely a lot on downsampling the data by a large factor should carefully balance the risk of losing several low S/N pulsars against gains in sensitivity for other types of populations of pulsars (for e.g. relativistic binary pulsars).

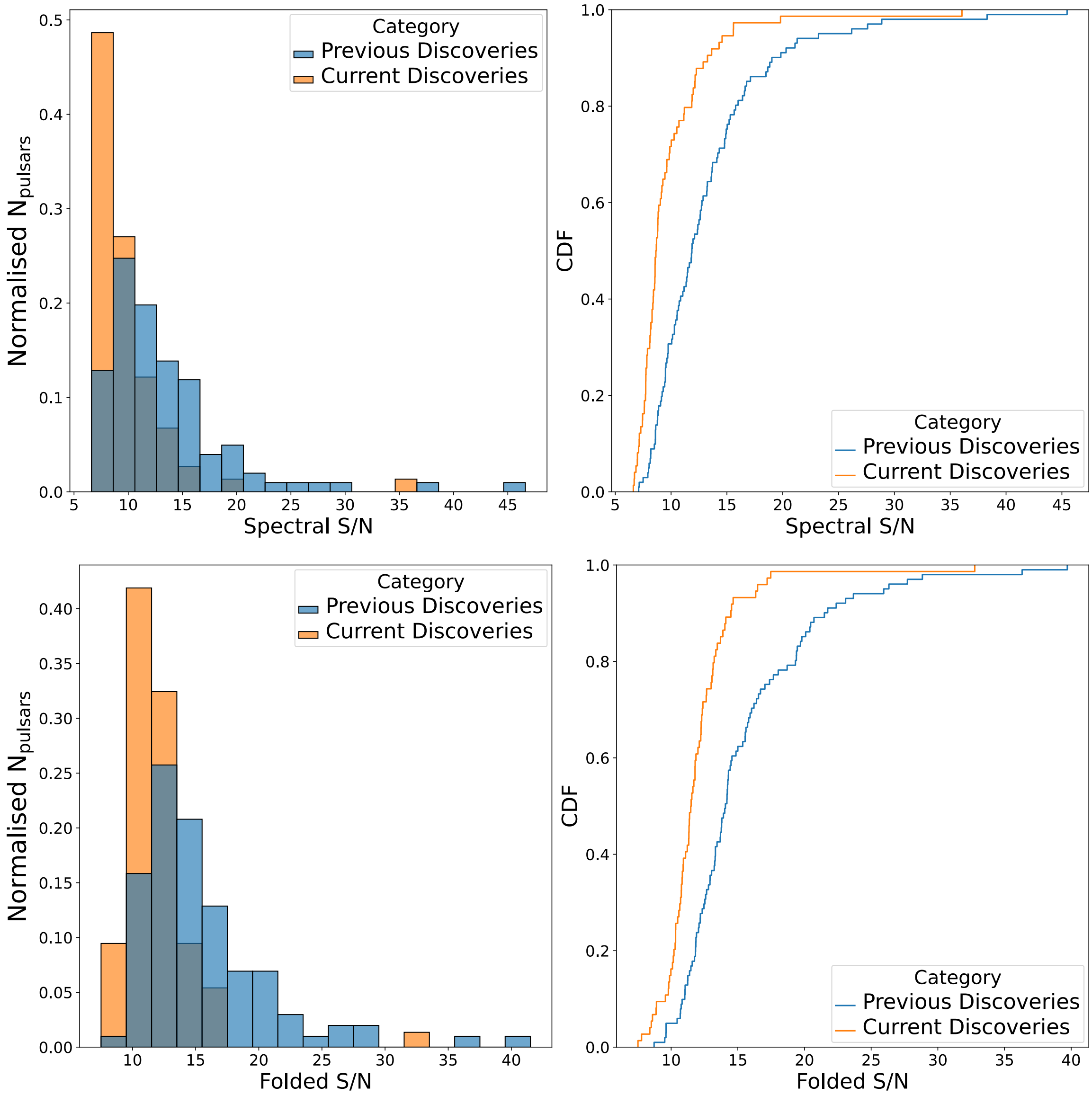


Figure 3.16: Histogram and Cumulative distribution function of S/N measurements of current discoveries compared to previous discoveries from LOWLAT. In the top, we show the distribution of the FFT spectral S/N and in the bottom we show the distribution of folded S/N. A large number of the new discoveries were made possible due to the lower FFT spectral S/N threshold which was feasible due to gains in computational power that enabled us to fold more pulsar candidates and find many relatively low S/N pulsars. find See main text for further discussion.

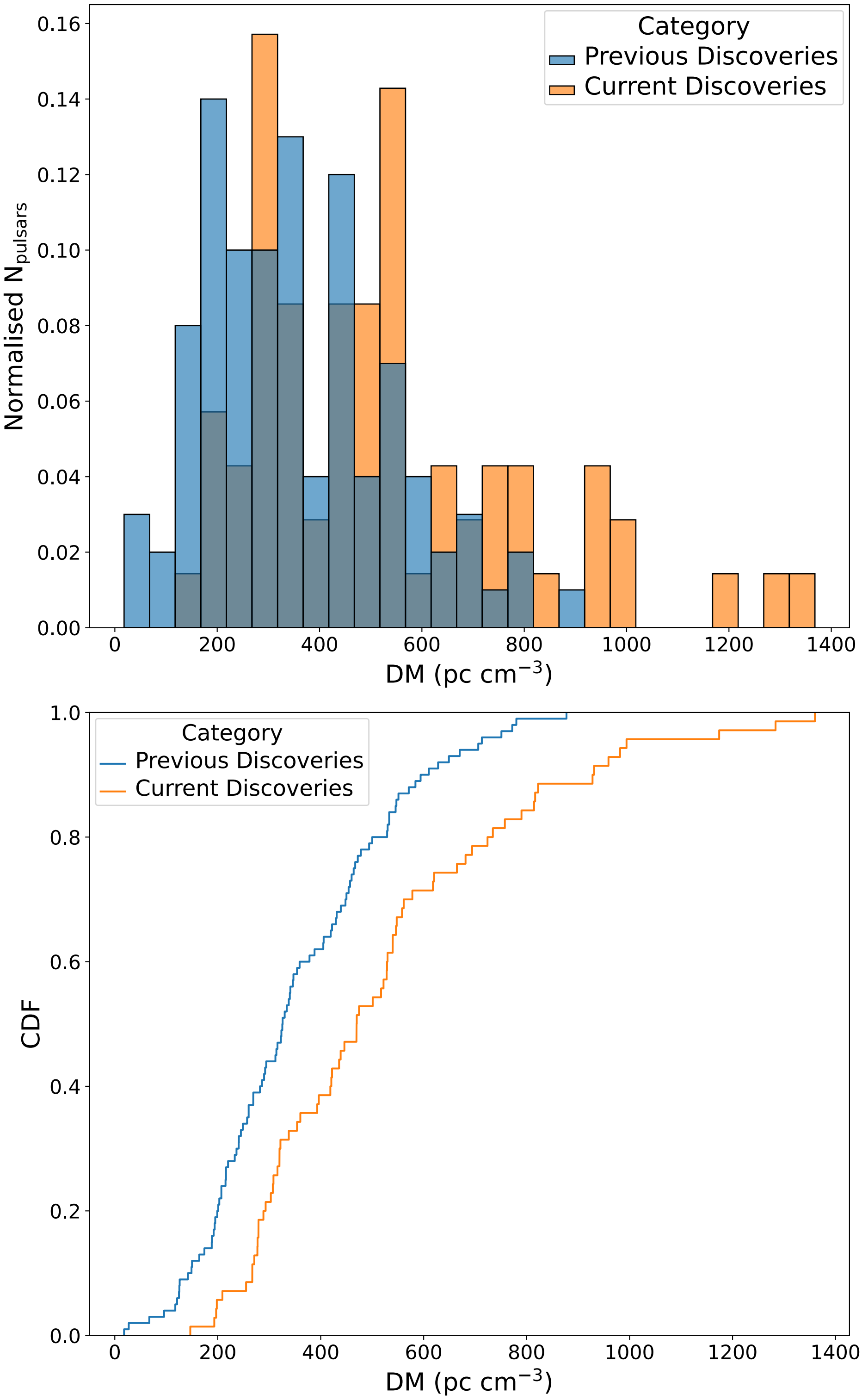


Figure 3.17: Detected DM Distribution of the pulsar discoveries presented in this work ('Current' Discoveries) to pulsar discoveries reported in past search efforts in LOWLAT (Ng et al., 2015; Cameron et al., 2020). In the top we show a histogram and in the bottom we show a CDF of the DM distribution for both populations. On average, pulsars found in the reprocessing are probing pulsars at a higher DM and therefore are likely located further away from us.

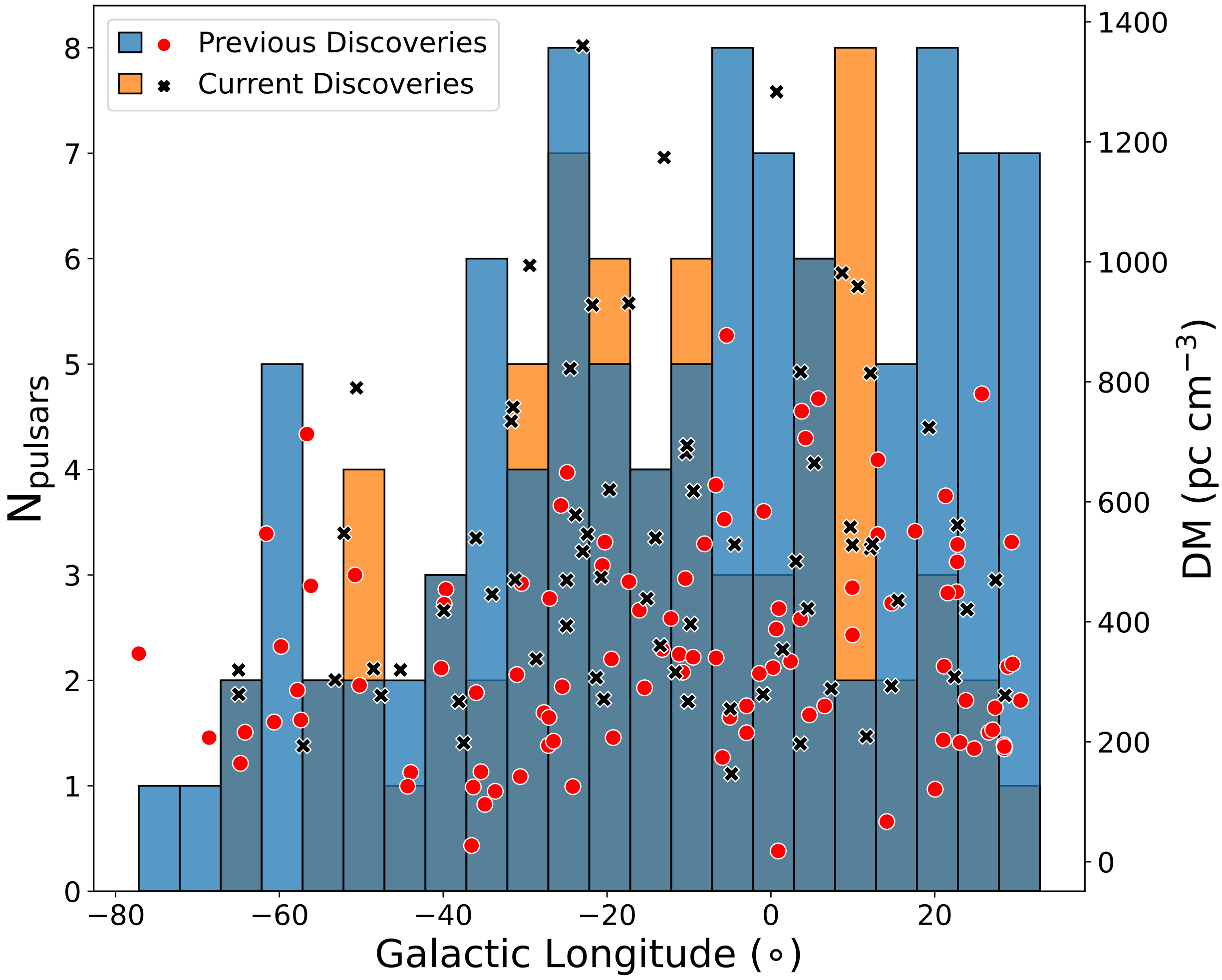


Figure 3.18: Histogram of Galactic longitude distribution of the current and previous discoveries in LOWLAT. The scatter plot markers show the detected DM of each pulsar. The Gl distribution of both populations are broadly consistent with each other. See main text for discussion.

Table 3.10: Redetection of 50 new pulsars found in the reprocessing of Lowlat using the template-bank pipeline. Out of this, 47 pulsars were discovered by ongoing reprocessing efforts using an acceleration search pipeline- peasoup (Sengar et al., prep). Three pulsars (marked in the footnotes) were discovered by the FFA pipeline (Wongphechauxsorn et al., prep). Similar to table 3.6, we give pulsars a temporary name with two digits of declination unless the pulsar has been redetected in the ongoing MGPS-L survey using Meerkat. We report here the barycentric spin period P, dispersion measure DM and barycentric period epoch PEPOCH. These values have been taken from the discovery observation from lowlat, and values in parenthesis represent our uncertainty in the final digit. We also report DM-dependent distances based on 'NE2001' model (Cordes & Lazio, 2002) and the YMW16 model (Yao et al., 2017)

PSRJ	pointing/beam	gl (°)	gb (°)	P (ms)	DM (pc cm^{-3})	S/N	PEPOCH MJD	$D_{\rm YMW16}$ (kpc)	$D_{\rm NE2001}$ (kpc)
J1136−64	2012-07-23-00:10:48/10	295.016	-2.967	1023.619(3)	309(9)	14.28	56131.033387	7.07	3.17
J1325−6256[b]	2011-12-10-16:54:46/10	306.798	-0.326	28.970674(2)	303.6(2)	14.2	55905.726955	6.39	5.38
J1333−61	2011-09-18-22:03:34/07	307.882	0.650	1532.134(6)	546(14)	16.79	55822.942569	8.97	12.02
J1348−62	2011-10-10-21:55:27/06	309.415	-0.350	616.2599(9)	685(7)	13.58	59179.818916	10.72	12.21
J1406−59	2011-04-19-15:27:52/07	312.415	2.081	1248.312(4)	286(12)	12.94	55670.673697	5.39	5.65
J1437−62	2011-12-05-20:04:47/06	314.772	-2.390	777.993(1)	323(7)	13.41	55900.858247	6.89	6.51
J1518−60	2012-07-31-12:35:11/03	320.068	-2.595	510.6555(7)	420(5)	14.91	56139.552401	8.71	15.15
J1521−57	2013-02-02-16:02:33/12	321.912	-0.574	173.68169(9)	270(2)	10.08	56325.692563	4.26	4.49
J1548−55	2011-12-29-17:55:52/06	325.972	-1.170	541.382(1)	446(11)	8.91	55924.768219	6.43	6.54
J1555−53	2011-05-17-11:18:22/03	328.529	-0.481	1170.895(9)	785(25)	11.86	55698.501771	9.90	6.42
J1600−49	2012-01-04-18:26:39/04	331.344	2.499	287.5260(4)	328(5)	10.60	55930.790073	9.38	6.90
J1603−54	2011-12-07-21:23:21/04	328.773	-1.150	960.792(3)	472(9)	12.76	55902.912297	7.26	6.35
J1632−49	2011-12-13-04:39:04/12	335.499	-0.735	416.836(2)	814(13)	11.93	55908.214809	9.10	7.07
J1634−49	2011-04-23-20:03:59/02	335.085	-1.571	356.7153(4)	465(3)	14.51	55674.865709	6.52	7.01
J1635−46	2011-06-26-16:09:12/11	337.596	0.449	1488.902(7)	549(14)	14.89	55738.704021	6.14	4.97
J1638−47	2012-04-14-13:41:06/02	337.035	-0.549	426.668(2)	1374(19)	12.35	56031.599597	17.36	10.09
J1639−46	2012-04-12-12:47:43/13	338.208	0.224	519.136(1)	917(7)	8.78	56029.561993	9.18	5.80
J1641−49	2011-10-13-06:08:35/09	336.154	-2.158	795.189(2)	578(9)	9.86	55847.279164	10.17	18.44
J1647−49	2011-10-10-05:13:04/06	337.037	-2.589	247.5278(2)	515(2)	11.73	55844.240911	9.77	18.90
J1651−42	2012-12-30-02:45:46/07	342.653	1.080	496.547(2)	943(14)	12.81	56291.135731	13.30	25.00
J1651−46	2011-12-07-22:36:35/04	339.273	-1.551	569.351(2)	487(11)	13.08	55902.962574	6.64	11.47
J1655−40	2011-04-20-20:48:56/03	344.871	2.069	276.6890(5)	439(6)	11.78	55671.896549	6.75	18.20
J1700−39	2011-12-10-04:41:41/01	345.910	1.480	3746.46(4)	509(36)	15.08	55905.215785	6.84	15.49
J1710−39	2011-12-29-20:22:33/01	346.980	0.060	977.337(8)	1198(29)	12.12	55924.869521	12.57	7.45
J1717−41	2011-10-13-07:22:02/04	346.487	-1.989	546.2329(8)	356(5)	16.45	55847.330481	5.53	11.82

continued in next page

PSRJ	pointing/beam	gl (°)	gb (°)	P (ms)	DM (pc cm^{-3})	S/N	PEPOCH MJD	$D_{\rm YMW16}$ (kpc)	$D_{\rm NE2001}$ (kpc)
J1719−36	2012-08-03-08:18:01/09	350.548	0.274	757.151(2)	624(7)	12.82	56142.375630	6.61	4.65
J1720−37[a]	2011-07-06-09:13:30/13	350.229	0.024	934.085(4)	388(13)	12.33	55748.415402	4.61	4.92
J1723−40	2012-11-25-01:22:55/13	348.360	-1.160	1982.26(1)	347(19)	8.42	56256.077862	4.59	4.44
J1735−28	2011-06-27-16:02:05/12	359.067	2.103	428.5601(7)	279(5)	9.97	55739.699884	4.44	8.03
J1738−33	2012-10-04-10:29:13/09	355.042	-1.385	357.7344(4)	257(3)	10.04	56204.461351	3.90	3.97
J1739−26	2011-06-30-16:47:36/07	1.441	2.503	490.1418(8)	354(5)	10.56	55742.729027	6.05	16.12
J1742−34[a]	2011-05-08-14:10:36/08	354.516	-2.591	1116.396(9)	157(21)	10.3	55689.621159	3.07	2.81
J1746−2829[a]	2011-05-17-15:33:25/02	0.451	0.113	1888.96(1)	1318(21)	10.74	55698.678994	15.80	8.23
J1757−26	2011-10-06-07:02:33/12	3.060	-1.124	354.458(3)	502(24)	11.77	55840.317833	7.02	9.44
J1758−24	2011-04-24-21:02:42/03	5.292	-0.160	633.105(2)	652(9)	12.86	55675.905772	10.41	4.55
J1758−25	2013-04-04-17:59:06/13	4.491	-0.774	605.501(1)	415(6)	9.45	56386.776920	5.48	4.23
J1806−21	2013-01-07-04:18:23/08	8.690	-0.328	328.558(1)	989(14)	12.25	56299.199796	12.22	6.25
J1808−14	2013-04-06-16:18:54/03	14.715	2.491	836.422(2)	307(8)	11.47	56388.706730	5.42	9.78
J1808−19	2011-07-02-12:25:09/03	10.625	0.261	101.6661(1)	969(3)	16.44	55744.549250	11.71	6.47
J1809−20	2013-04-02-16:32:16/06	9.952	-0.570	57.25620(2)	528(1)	11.44	56384.716057	6.65	4.50
J1813−14	2013-04-02-21:22:53/04	15.525	1.480	1035.407(3)	427(10)	13.91	56384.917350	6.11	11.87
J1814−1845[b]	2013-04-08-18:54:20/05	12.139	-0.709	1089.989(4)	534(10)	33.6	56390.815126	4.89	7.07
J1814−18	2013-04-08-18:54:20/05	12.157	-0.740	684.209(3)	824(16)	10.62	56390.815133	13.42	11.90
J1817−19	2013-04-08-18:54:20/10	11.660	-1.583	1229.084(4)	191(12)	12.11	56390.815078	3.69	3.60
J1823−11	2011-12-31-00:27:33/06	19.295	0.659	286.184(1)	725(15)	10.39	55926.039287	8.34	6.99
J1830−09	2011-07-05-14:23:13/09	22.430	0.424	695.490(1)	309(7)	15.26	55747.631059	4.62	4.09
J1835−09	2013-04-01-18:01:40/11	22.780	-0.510	750.461(5)	555(22)	12.54	56294.039346	6.66	4.94
J1844−09	2013-04-01-20:26:44/01	23.950	-2.580	634.445(1)	415(6)	11.86	56383.877573	8.10	18.09
J1846−05	2012-04-13-20:25:11/06	27.453	-1.350	1444.984(6)	474(14)	14.89	56030.877942	7.75	9.07
J1854−05	2011-12-13-06:05:17/01	28.610	-2.980	1279.944(5)	275(12)	12.60	55908.274585	5.94	8.24

[a] These pulsars were initially discovered by the ongoing reprocessing of LOWLAT using an FFA pipeline (Wongphechauxsorn et al., prep).

[b] Positions reported for these pulsars were obtained from the ongoing MGPS-L survey (Padmanabh et al., prep) but the Period, DM and S/N reported in these rows are from discovery Parkes Observation.

3.3.7 Status of Survey Processing and Evaluation of survey yield

We have completed searching 14,488 beams in Lowlat out of a total of 16,637 which comprises of 87.1% of the total survey. In figure 3.19, we show the spatial configuration of HTRU pointings processed during this thesis. Green circles indicate beams processed and red circles indicate pointings that are still to be processed. Given the extensive number of new pulsars that have been presented in this thesis along with a list of the previously missed known pulsars shown in table A.1, it is now possible to make a revised evaluation of the survey yield. Keith et al. (2010) estimated that LOWLAT would find '957' normal pulsars (defined as $P > 30\,\mathrm{ms}$) using simulations from the PSRPOP[30] software package. Ng et al. (2015) later revised this estimate to 1020 normal pulsars based on updated simulations using the software PSRPOPPY[31] (Bates et al., 2014). Both these results are statistically consistent, and we will use the latter to stay consistent with past discussions of survey yield presented in Ng et al. (2015); Cameron et al. (2020). The first-pass pipeline was used to process 94.1% of the survey. Remaining portions of the survey were left our due to difficulty in accessing archival storage tapes. Therefore, re-scaling the total predictions to the part of the data that was processed should yield 960 normal pulsars. Cameron et al. (2020) reported a total of 723 normal pulsar detections combining new pulsars presented in that work with known pulsar redetections and discoveries presented in Ng et al. (2015), and reported falling short of predictions from simulations by $\sim$ 25%. As mentioned before in section 3.3.6, we know that this reduced discovery rate was caused due to incorrect bit rescaling while downsampling, high DM tolerance for dedispersion steps, high FFT S/N and candidate rank threshold used before folding as well as human error. In this work, we presented 20 new pulsars, redetection of PSR J1753−2819 along with a redetection of 50 new pulsars and an additional 60 known pulsars which were missed in the original pipeline. Some of our known pulsar redetections include the newly discovered 2.772 ms MSP J1431-6331 from ASKAP (Kaplan et al., 2019), a 1.492 ms MSP J1804−2858 recently found in the reprocessing of Med-lat (Morello et al., 2019), PSR J1809−2036 which was detected initially in the High-Lat pulsar survey (Lina Levin, 2019, HitRun mailing list comm) and confirmed in LOWLAT. A further 22 pulsars are expected to be announced in Sengar et al. (prep). Combining all these, along with the initial list provided by Cameron et al. (2020) takes the total to 872 unique normal pulsar detections which is 9% short of predictions from PSRPOPPY. An additional 7 % of the data still remains to be processed in order to be consistent with the amount of data that was processed by the first-pass pipeline. Therefore, extrapolating our numbers to this leads to 884 total pulsar detections. The simulations done using PSRPOPPY assumed a detection S/N of 9.0, whereas we inspected candidates below this limit which was made possible due to our machine learning search pipeline (Balakrishnan et al., 2021). Out of our total pulsar detections, at-least 10 pulsars have a S/N below 9, therefore subtracting this from our total detections in order to stay consistent with the simulation threshold, we get 874 normal pulsar detections which falls short by $\sim$ 9%. This can be partially

---

[30] `http://psrpop.sourceforge.net`

[31] `https://github.com/samb8s/PsrPopPy`

explained by the fact that searches on the full native time resolution of LOWLAT is still ongoing, and the extra known pulsars reported here were done based on searches on decimated data. Additionally, propagation effects due to scintillation, and scattering could also play a major role in reducing the number of our total detections. Also, PSRPOPPY assumes sensitivity of a pulsar survey based on the radiometer equation, however Lazarus et al. (2015) reported that the true sensitivity of a pulsar survey is a function of period, DM, duty cycle of the pulsar, the RFI environment and slow pulsars are adversely affected due to the influence of red noise. The true sensitivity of LOWLAT could be evaluated by injecting signals into the data using the technique described in Lazarus et al. (2015) and revising the expected discoveries based on that. Additionally, both Keith et al. (2010) and Ng et al. (2015) do not report a standard deviation in the expected number of pulsar discoveries. These predictions were based on the 'snapshot' method (Bates et al., 2014) where we populate the Galaxy with pulsars with a certain luminosity, spin-period and spatial distribution till a given standard pulsar survey (in this case PMPS) detects the total number of pulsars (1038) it actually found. Once this is done based on telescope, receiver parameters and sky coverage of future pulsar surveys, it makes a prediction on how many new pulsars will be found in that survey. However, we know that the total number of PMPS discoveries gets updated over time with additional reprocessing with more sophisticated algorithms. An additional 20 new pulsar discoveries (besides the original 1038 pulsar detections) from the latest leg of PMPS reprocessing are expected to be announced soon (R.Sengar, 2021, priv. comm.). Also, PSRPOPPY assume a log-normal luminosity distribution for normal pulsars from Lorimer et al. (2006) which is not well constrained at the lower luminosity end. Therefore, any predictions from PSRPOPPY should be evaluated with caution. The ongoing processing of LOWLAT using the FFA pipeline could also lead to the discovery of several new normal pulsars as FFA is a more optimal approach to find normal pulsars with narrow duty cycle. Separate predictions were made about the total amount of MSP discoveries in LOWLAT. This is because MSPs tend to have different selection effects in a pulsar survey compared to normal pulsars and additionally the number of known Galactic MSPs is much lesser than the number of known normal pulsars, therefore, making future discovery predictions more challenging. Keith et al. (2010) estimated that LOWLAT would find 51 MSPs which was revised to 43 by Ng et al. (2015). Levin et al. (2013) predicted an increased yield of 68 MSPs. Using these predictions as lower and upper limits, and rescaling it to 94.1% of the survey which was processed by the first-pass pipeline Cameron et al. (2020) reported that we should expect to find 40-64 MSP detections whereas only 18 MSPs were found with the first-pass pipeline. The total number of MSP detections so far from the reprocessing pipeline for 87.1 % of the survey is 34, almost double the number initially reported. A large part of these new discoveries were made possible due to FFT searches on the full native time resolution observations of LOWLAT (Sengar et al., prep). Extrapolating these numbers to 94.1% should lead to 38 MSP detections. None of our confirmed MSP detections have a folded S/N below 9, therefore these numbers are consistent with the detection threshold used in PSRPOPPY. As mentioned earlier, it is likely that our pipeline would have missed multiple isolated MSPs rotating faster than spin-period of 13 ms due to us targeting PSR-BH binaries which requires investing more compute time for orbital searches. We expect more known MSPs to be detect by a pure FFT based search on the full-length observations at 64 µs sampling. Overall, once all pipelines have finished processing we expect the actual number of MSPs found in LOWLAT should match the lower limit of predictions made from simulations.

3.3.8 Flux-Density Limits on Short Orbital Period PSR-BH Binaries

Using the template-bank algorithm, we have conducted the most sensitive search for PSR-BH binary systems in the Southern Galactic-Plane at L-band. In the following sections, we will quantify this extra sensitivity from our pipeline and use the null results from our searches to place limits on the detectable Population of PSR-BH binaries in the Galaxy. The HTRU South LOWLAT survey as a whole has now been searched by at least three independent pipelines. First was a Sigproc-based Pipeline implementing a segmented search (here after called segmented accel-search) described in Ng et al. (2015). Second is a GPU-based acceleration search pipeline (here after called GPU accel-search) that expanded the maximum acceleration value in the full length 72-minute observation of LOWLAT and the discoveries from this is expected to be announced in the near future (Sengar et al., prep). Finally, we have the template-bank pipeline which is the subject of this chapter that focused on the phase-space of PSR+Stellar Mass Black holes where 10-40% orbit is visible in the full length observation. The maximum acceleration value chosen for the accel.search and template-bank pipeline where based on the values expected for PSR-BH binaries in circular orbits. Given the null results for all three, we can now use the known sensitivity of each pipeline to place a flux density limit on PSR-BH binaries for various orbital parameters. Ng et al. (2015) reported that the minimum flux density (also known as sensitivity) for the HTRU-S Lowlat survey is 0.1 mJy. The true sensitivity of a survey is not a fixed value but is a function of spin-period of the pulsar, duty cycle, DM, and the RFI environment (see Lazarus et al. 2015 for a more detailed discussion on this). In this section for simplicity, we assume a fixed value as our best-case minimum detectable flux density which is generally valid for isolated pulsars. While searching for binary pulsars, if the orbit of the pulsar is not known our sensitivity decreases due to the apparent change in the observed period of the pulsar due to Doppler modulation. Johnston & Kulkarni (1991); Bagchi et al. (2013) described an analytical formulation for quantifying the S/N regained by applying a binary search pipeline called efficiency factors assuming standard FFT search, an acceleration search and a jerk search. We expand on this work by first simulating observations of a PSR-BH binary in circular orbits with observational parameters resembling a typical full-length LOWLAT observation (see table 3.4). We then run all three pipelines i.e the segmented accel-search, the GPU accel. search and the template-bank pipeline. For the segmented accel-search pipeline, we divide the observation into halves, quarters and eighths and search with the same acceleration range reported in Ng et al. (2015). We then take the best case S/N and use that for our minimum detectable flux density estimate. For the GPU accel. search pipeline, we use a maximum acceleration value of 100 m s^{-2} on the full-length. For the template-bank

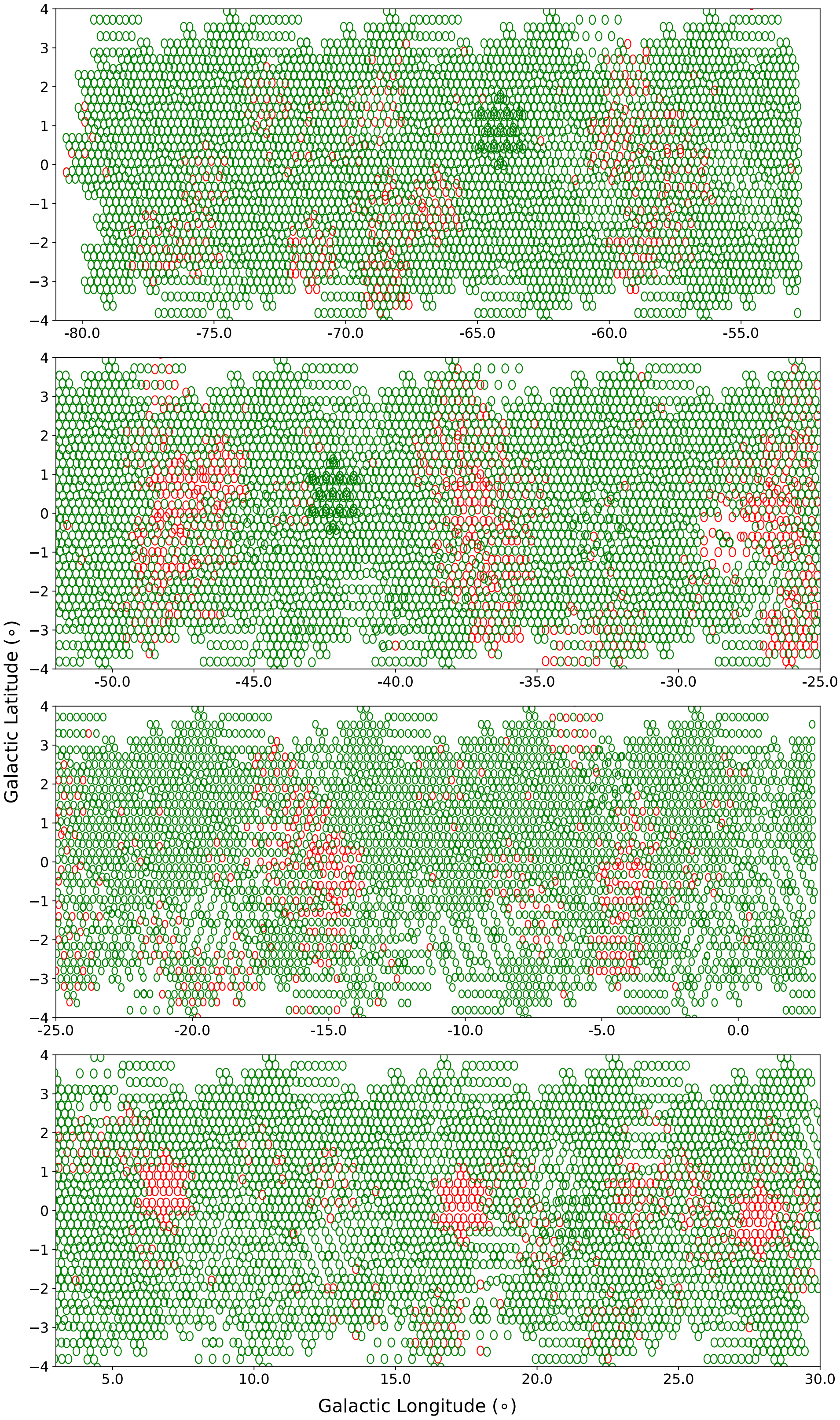


Figure 3.19: Spatial distribution of HTRU pointings on the sky. Green circles indicate pointings processed during our reprocessing. Red circles indicate pointings that are still to be processed.

pipeline, we run both our searches, and report here the best case minimum detectable flux density estimate. We start by fixing the spin-period of the pulsar $P_{\rm spin} = 10\,{\rm ms}$, duty cycle = 20 % and varying the orbital period of the binary from 4 hours to 24 hours and the mass of the black hole from 5 $M_\odot$ to 15 $M_\odot$. We use an inclination of 60° for our simulations, pulsar mass of 1.4 $M_\odot$ and we let the initial orbital phase vary randomly between 0 and 2 π radians. We then calculate the minimum detectable flux density $S_{\rm min}$ for each pipeline with our theoretical minimum detectable flux density being 0.1 mJy (Ng et al., 2015). Our results are shown in figures 3.20 and 3.21. As expected, the GPU accel. search pipeline (bottom of figure 3.20) is sensitive to long orbital period binaries above 18 hours with the minimum flux density estimate ranging between 0.1 to 0.2 mJy. A minimum flux density estimate of 0.2 mJy is a 50 % degradation from the theoretical maximum sensitivity of an isolated pulsar. As the orbit gets more compact ($P_{\rm orb}$ = 8-18 hrs), our sensitivity range drops to between 0.2-0.43 mJy which in the worst case is a 76.7% loss in sensitivity. For orbits with shorter orbital periods ($P_{\rm orb}$ = 4-8 hrs), our sensitivity range is between 0.4-1.4 mJy which in the worst case is a 92.8 % loss in sensitivity. The general trend is that as the orbit gets more compact and the mass of the black hole increases, our search sensitivity drops. The slight non-uniformity in sensitivity in adjacent cells for orbits with similar orbital period and companion mass is due to the fact that we have let the orbital phase of the binary vary between 0 and 2 π radians[32]. Cells at the same grid location for each pipeline however are directly comparable. In the top part of figure 3.20, we show the results from the segmented accel-search pipeline (Ng et al., 2015). In section 3.3.6, we discussed some of the drawbacks of the segmented search pipeline which limited the maximum acceleration range in the full 72-minute observation to $a_{\rm max} = \left|1\,{\rm m\,s^{-2}}\right|$, here, we see some of its advantages. The segmented-search offers increased sensitivity throughout the binary phase space of PSR-BH binaries with minimum detectable flux density ranging between 0.14-0.37 mJy which in the worst case is 73 % loss. In the top part of figure 3.21, we show the results from our template-bank pipeline. For orbits longer than $P_{\rm orb} > 20$ hrs, the performance is comparable with the segmented acceleration search pipeline and for some cases it performs worse because those orbits are outside our search range. However, as the orbit gets more compact, this is when higher order searches like the template-bank pipeline truly shine. We obtained an increased sensitivity towards orbits in the $P_{\rm orb}$ = 6-12 hrs regime which was the range chosen in our template-bank. Our sensitivity here ranges between 0.1-0.2 with a median sensitivity of 0.114 mJy which is an average loss of 12.2% and in the worst case a loss of 50 %. This can be better visualised in the bottom part of figure 3.21 where we show the sensitivity improvement factor from the template-bank pipeline compared to the segmented accel-search pipeline. A factor above unity is an improvement for the template-bank pipeline and conversely below unity indicates better performance from the segmented accel. search pipeline. We have at least doubled the sensitivity or in other words halved the minimum detectable flux density for PSR-BH binaries for orbits in the 6-12 hour

---

[32] A favorable orbital phase can make an otherwise difficult to detect relativistic binary pulsar more easily detectable.

range compared to past searches in LOWLAT. A factor of 2 improvement purely from a change in search algorithm is quite significant as from the radiometer equation we know that to produce a similar sensitivity gain for isolated pulsars would require four times the original observation length. We would like to emphasize that such an improvement from acceleration-search cannot be obtained irrespective of the maximum acceleration value used in the search due to the previously defined 10% rule. It is clear from figure 3.21 that we have now probed this compact orbit phase-space with a much more stringent flux density limit that was previously not possible. Naturally, our minimum detectable flux density limit increases for orbits shorter than 6 hours due to these ranges being outside our template-bank and is much worse than the segmented acceleration search pipeline. We leave these parameter spaces to be explored by future reprocessing efforts including the possibility of an application of a segmented template-bank search pipeline which can provide more stringent limits that the ones currently available in this orbital period regime. We also repeated these simulations for a pulsar with a spin-period 100 ms and 1000 ms. These results are shown in Appendix B. It is in general easier to find normal pulsars orbiting black holes compared to MSPs. Therefore, in order to see similar gains in sensitivity its important to ensure that the template-bank is designed for progressively shorter orbits or more massive black holes as was the case for our Non-Recycled Pulsar-BH binary search described in section 3.2.4.

3.3.9 Limits on the Detectable Fraction of PSR-BH Binaries with short orbits

In this section, we do a similar sensitivity comparison of each pipeline by computing the minimum detectable flux density as done in the previous section but in the orbital-period and pulsar spin-period plane keeping the mass of the black hole fixed at $8\mathrm{M}_{\odot}$. We restrict our discussion to binaries with an orbital period range of $P_{\mathrm{orb}} = 4 - 24$hrs and a pulsar spin-period between 5 to 100ms. As is the case in the previous section, we use an inclination of 60° for our simulations, pulsar mass of 1.4 $\mathrm{M}_{\odot}$ and we let the initial orbital phase vary randomly between 0 and 2 π radians. Here, for brevity we only show the results from the template-bank pipeline in figure 3.22. Similar plots assuming a segmented acceleration search pipeline and the GPU accel-search pipeline including the gain factor defined in the previous section can be found in Appendix C.1 and C.2 respectively. As is the case for our black hole mass vs orbital period plot, we have a uniform sensitivity towards PSR-BH binaries for parameter regions covered by our template-bank. The slight non-uniformity in sensitivity (for e.g at Porb = 13.63, 14.98 and spin-period=5.85, 6.85 respectively) are due to probabilistic holes in our template-bank. We have assumed a coverage of $\eta = 0.9$ and a mismatch $m = 0.2$, therefore, statistically we are likely to have small regions of reduced sensitivity. Once, we calculate sensitivity we then assign a luminosity to each pulsar in this plane by sampling from a log-normal luminosity distribution[33] (Faucher-Giguère & Kaspi, 2006)

---

[33] We have ignored geometric effects here like beaming fraction. Therefore such luminosity distribution is referred to in literature as “pseudo-luminosity” distribution (Arzoumanian et al., 2002; Lorimer, 2008).

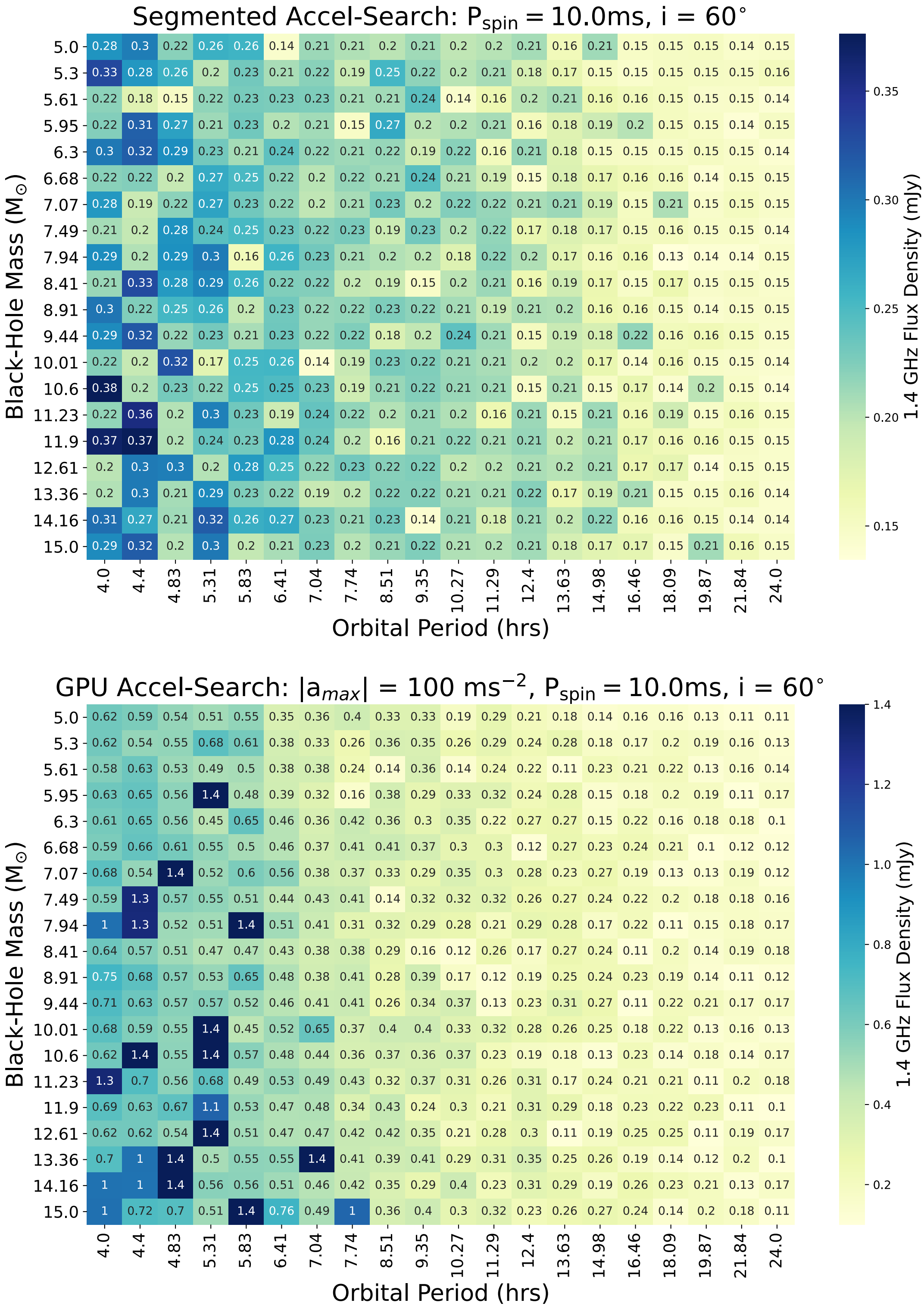
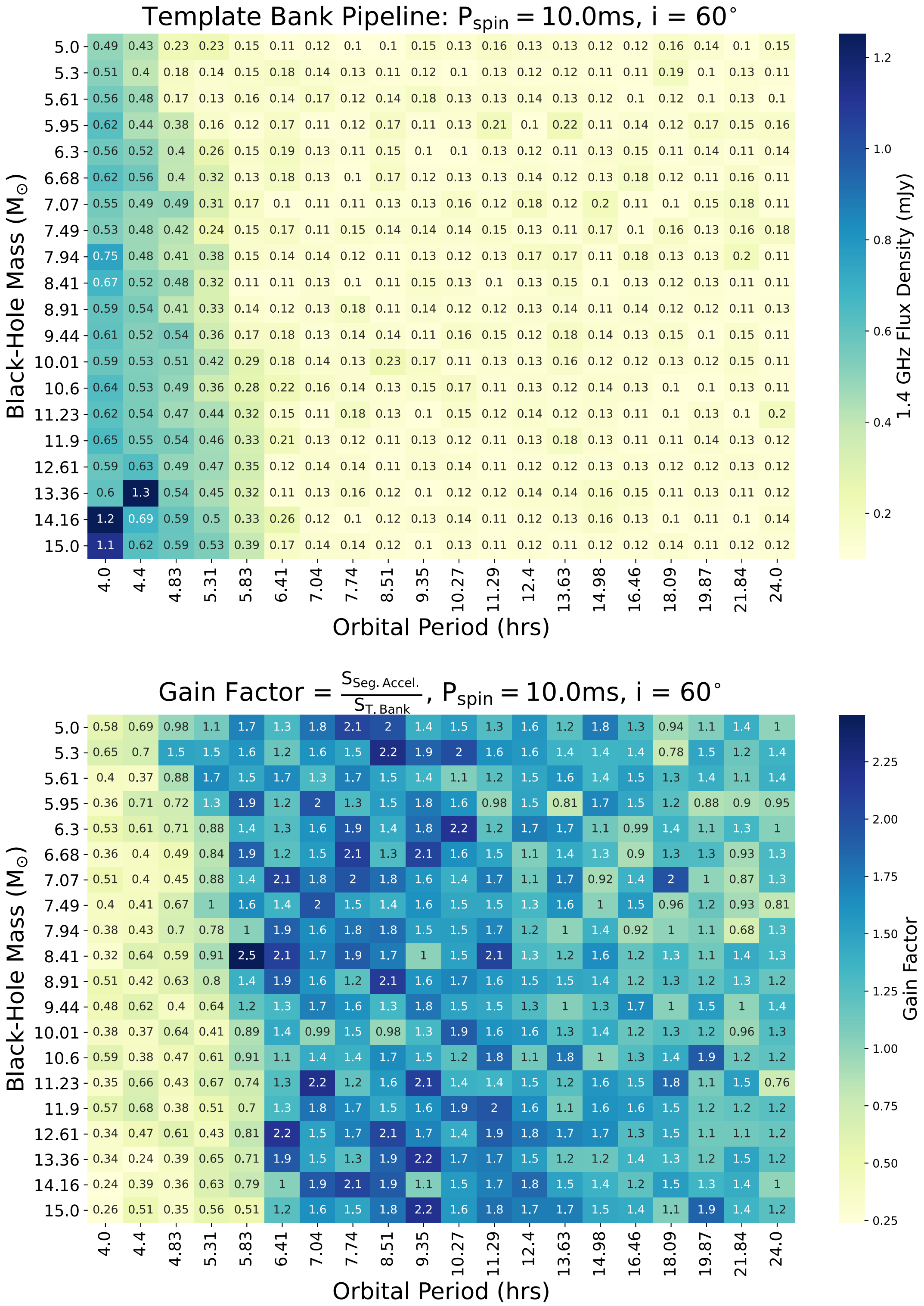


Figure 3.20: Minimum detectable flux density limit for a 10 ms pulsar orbiting a black hole of varying mass as a function of the orbital period of the binary. In the top we show the results from a segmented acceleration search pipeline (Ng et al., 2015) and in the bottom we show the results from the ongoing reprocessing of LOWLAT using a GPU acceleration search (Sengar et al., prep). See text for full discussion.



Template Bank Pipeline: $P_{spin} = 10.0$ms, i = 60°

	4.0	4.4	4.83	5.31	5.83	6.41	7.04	7.74	8.51	9.35	10.27	11.29	12.4	13.63	14.98	16.46	18.09	19.87	21.84	24.0
5.0	0.49	0.43	0.23	0.23	0.15	0.11	0.12	0.1	0.1	0.15	0.13	0.16	0.13	0.13	0.12	0.12	0.16	0.14	0.1	0.15
5.3	0.51	0.4	0.18	0.14	0.15	0.18	0.14	0.13	0.11	0.12	0.1	0.13	0.12	0.12	0.11	0.11	0.19	0.1	0.13	0.11
5.61	0.56	0.48	0.17	0.13	0.16	0.14	0.17	0.12	0.14	0.18	0.13	0.13	0.14	0.13	0.12	0.1	0.12	0.1	0.13	0.1
5.95	0.62	0.44	0.38	0.16	0.12	0.17	0.11	0.12	0.17	0.11	0.13	0.21	0.1	0.22	0.11	0.14	0.12	0.17	0.15	0.16
6.3	0.56	0.52	0.4	0.26	0.15	0.19	0.13	0.11	0.15	0.1	0.1	0.13	0.12	0.11	0.13	0.15	0.11	0.14	0.11	0.14
6.68	0.62	0.56	0.4	0.32	0.13	0.18	0.13	0.1	0.17	0.12	0.13	0.13	0.14	0.12	0.13	0.18	0.12	0.11	0.16	0.11
7.07	0.55	0.49	0.49	0.31	0.17	0.1	0.11	0.11	0.13	0.13	0.16	0.12	0.18	0.12	0.2	0.11	0.1	0.15	0.18	0.11
7.49	0.53	0.48	0.42	0.24	0.15	0.17	0.11	0.15	0.14	0.14	0.14	0.15	0.13	0.11	0.17	0.1	0.16	0.13	0.16	0.18
7.94	0.75	0.48	0.41	0.38	0.15	0.14	0.14	0.12	0.11	0.13	0.12	0.13	0.17	0.17	0.11	0.18	0.13	0.13	0.2	0.11
8.41	0.67	0.52	0.48	0.32	0.11	0.11	0.13	0.1	0.11	0.15	0.13	0.1	0.13	0.15	0.1	0.13	0.12	0.13	0.11	0.11
8.91	0.59	0.54	0.41	0.33	0.14	0.12	0.13	0.18	0.11	0.14	0.12	0.12	0.13	0.14	0.11	0.14	0.12	0.12	0.11	0.13
9.44	0.61	0.52	0.54	0.36	0.17	0.18	0.13	0.14	0.14	0.11	0.16	0.15	0.12	0.18	0.14	0.13	0.15	0.1	0.15	0.11
10.01	0.59	0.53	0.51	0.42	0.29	0.18	0.14	0.13	0.23	0.17	0.11	0.13	0.13	0.16	0.12	0.12	0.13	0.12	0.15	0.11
10.6	0.64	0.53	0.49	0.36	0.28	0.22	0.16	0.14	0.13	0.15	0.17	0.11	0.13	0.12	0.14	0.13	0.1	0.1	0.13	0.11
11.23	0.62	0.54	0.47	0.44	0.32	0.15	0.11	0.18	0.13	0.1	0.15	0.12	0.14	0.12	0.13	0.11	0.1	0.13	0.1	0.2
11.9	0.65	0.55	0.54	0.46	0.33	0.21	0.13	0.12	0.11	0.13	0.12	0.11	0.13	0.18	0.13	0.11	0.11	0.14	0.13	0.12
12.61	0.59	0.63	0.49	0.47	0.35	0.12	0.14	0.14	0.11	0.13	0.14	0.11	0.12	0.12	0.13	0.13	0.12	0.12	0.13	0.12
13.36	0.6	1.3	0.54	0.45	0.32	0.11	0.13	0.16	0.12	0.1	0.12	0.12	0.14	0.14	0.16	0.15	0.11	0.13	0.11	0.12
14.16	1.2	0.69	0.59	0.5	0.33	0.26	0.12	0.1	0.12	0.13	0.14	0.11	0.12	0.13	0.16	0.13	0.1	0.11	0.1	0.14
15.0	1.1	0.62	0.59	0.53	0.39	0.17	0.14	0.14	0.12	0.1	0.13	0.11	0.12	0.11	0.12	0.12	0.14	0.11	0.12	0.12

Gain Factor = $\frac{S_{Seg. Accel.}}{S_{T. Bank}}$, $P_{spin} = 10.0$ms, i = 60°

	4.0	4.4	4.83	5.31	5.83	6.41	7.04	7.74	8.51	9.35	10.27	11.29	12.4	13.63	14.98	16.46	18.09	19.87	21.84	24.0
5.0	0.58	0.69	0.98	1.1	1.7	1.3	1.8	2.1	2	1.4	1.5	1.3	1.6	1.2	1.8	1.3	0.94	1.1	1.4	1
5.3	0.65	0.7	1.5	1.5	1.6	1.2	1.6	1.5	2.2	1.9	2	1.6	1.6	1.4	1.4	1.4	0.78	1.5	1.2	1.4
5.61	0.4	0.37	0.88	1.7	1.5	1.7	1.3	1.7	1.5	1.4	1.1	1.2	1.5	1.6	1.4	1.5	1.3	1.4	1.1	1.4
5.95	0.36	0.71	0.72	1.3	1.9	1.2	2	1.3	1.5	1.8	1.6	0.98	1.5	0.81	1.7	1.5	1.2	0.88	0.9	0.95
6.3	0.53	0.61	0.71	0.88	1.4	1.3	1.6	1.9	1.4	1.8	2.2	1.2	1.7	1.7	1.1	0.99	1.4	1.1	1.3	1
6.68	0.36	0.4	0.49	0.84	1.9	1.2	1.5	2.1	1.3	2.1	1.6	1.5	1.1	1.4	1.3	0.9	1.3	1.3	0.93	1.3
7.07	0.51	0.4	0.45	0.88	1.4	2.1	1.8	2	1.8	1.6	1.4	1.7	1.1	1.7	0.92	1.4	2	1	0.87	1.3
7.49	0.4	0.41	0.67	1	1.6	1.4	2	1.5	1.4	1.6	1.5	1.5	1.3	1.6	1	1.5	0.96	1.2	0.93	0.81
7.94	0.38	0.43	0.7	0.78	1	1.9	1.6	1.8	1.8	1.5	1.5	1.7	1.2	1	1.4	0.92	1	1.1	0.68	1.3
8.41	0.32	0.64	0.59	0.91	2.5	2.1	1.7	1.9	1.7	1	1.5	2.1	1.3	1.2	1.6	1.2	1.3	1.1	1.4	1.3
8.91	0.51	0.42	0.63	0.8	1.4	1.9	1.6	1.2	2.1	1.6	1.7	1.6	1.5	1.5	1.4	1.2	1.3	1.2	1.3	1.2
9.44	0.48	0.62	0.4	0.64	1.2	1.3	1.7	1.6	1.3	1.8	1.5	1.5	1.3	1	1.3	1.7	1	1.5	1	1.4
10.01	0.38	0.37	0.64	0.41	0.89	1.4	0.99	1.5	0.98	1.3	1.9	1.6	1.6	1.3	1.4	1.2	1.3	1.2	0.96	1.3
10.6	0.59	0.38	0.47	0.61	0.91	1.1	1.4	1.4	1.7	1.5	1.2	1.8	1.1	1.8	1	1.3	1.4	1.9	1.2	1.2
11.23	0.35	0.66	0.43	0.67	0.74	1.3	2.2	1.2	1.6	2.1	1.4	1.4	1.5	1.2	1.6	1.5	1.8	1.1	1.5	0.76
11.9	0.57	0.68	0.38	0.51	0.7	1.3	1.8	1.7	1.5	1.6	1.9	2	1.6	1.1	1.6	1.6	1.5	1.2	1.2	1.2
12.61	0.34	0.47	0.61	0.43	0.81	2.2	1.5	1.7	2.1	1.7	1.4	1.9	1.8	1.7	1.7	1.3	1.5	1.1	1.1	1.2
13.36	0.34	0.24	0.39	0.65	0.71	1.9	1.5	1.3	1.9	2.2	1.7	1.7	1.5	1.2	1.2	1.4	1.3	1.2	1.5	1.2
14.16	0.24	0.39	0.36	0.63	0.79	1	1.9	2.1	1.9	1.1	1.5	1.7	1.8	1.5	1.4	1.2	1.5	1.3	1.4	1
15.0	0.26	0.51	0.35	0.56	0.51	1.2	1.6	1.5	1.8	2.2	1.6	1.8	1.7	1.7	1.5	1.4	1.1	1.9	1.4	1.2

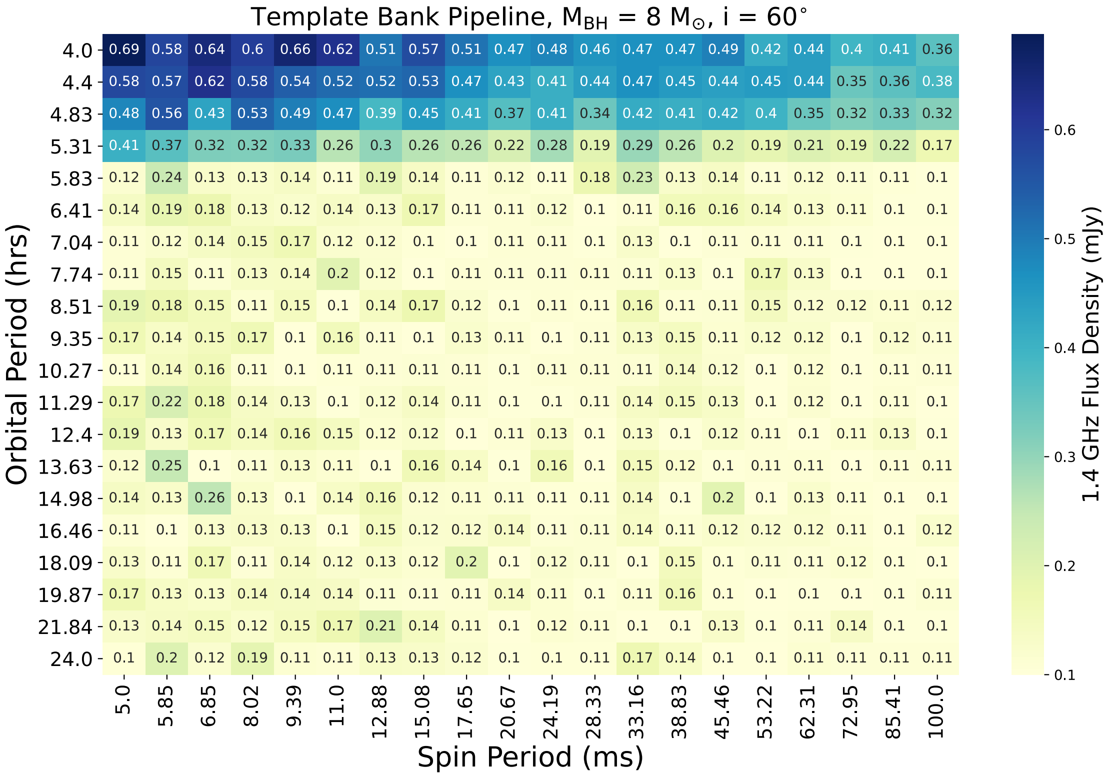
Figure 3.21: Minimum detectable flux density limit for a 10 ms pulsar orbiting a black hole of varying mass as a function of the orbital period of the binary. In the top we show the results from the template-bank pipeline which is presented in this work and in the bottom we show the sensitivity gain factor compared to the best-case acceleration-search pipeline i.e the segmented acceleration search pipeline (Ng et al., 2015). A factor above 1 indicates improvement from the template-bank pipeline and conversely below 1 indicates better performance from the acceleration-search pipeline. See text for full discussion.

with mean $\mu_{\log_{10}\mathrm{L}} = -1.1$ and standard deviation $\sigma_{\log_{10}\mathrm{L}} = 0.9$. In order to quantify a detectability fraction towards PSR-BH binaries in this plane, we also need to assume a certain distance between us and the binary pulsar. Depending on the part of the sky survey surveyed, and the detected DM of the binary, the distance to the PSR-BH can vary between a wide range of values. For simplicity, we first assume an isotropic distribution of PSR-BH binaries at a fixed distance of 6 kpc which is the median DM distance assuming the NE2001 electron density model of all the new pulsar discoveries found in LOWLAT as reported in Ng et al. (2015); Cameron et al. (2020) and this work. Using the luminosity and distance, we can compute the apparent flux density for each pulsar in this distribution and determine what fraction of sources lies above our 1.4 GHz minimum detectable flux density limit for each orbital and spin-period. These values are shown in figure 3.23. Our detectability fraction ranges between 2.6 to 3.6 % for binary orbits covered within our template-bank (i.e $\mathrm{P_{spin}} \geq 13$ms and $\mathrm{P_{orb}} = 6 - 12$hrs). 3.6 % is our best-case detectable fraction assuming an isolated pulsar and the theoretical sensitivity of LOWLAT. Our detectability range drops further between 1.3 % and 2.7 % for binaries with orbital period $P_{\mathrm{orb}} \leq 5.31$hrs due to losses from Doppler modulation as these orbits are not covered in our template-bank. These can in principle be improved by future reprocessing efforts. A large fraction of the binary pulsars assuming the luminosity distribution of Faucher-Giguère & Kaspi (2006) at this distance are simply too faint to be detected in LOWLAT. These numbers are a conservative upper-limit as we have assumed no beaming-fraction[34] losses in our analysis. Therefore, at this distance, we are largely sensitivity limited and future pulsar surveys using MeerKAT and the SKA[35] (Kramer, 2004) will be crucial to probe PSR-BH binaries in this regime. In figure 3.24, we marginalise over orbital period and spin-period and show a cumulative distribution function of the detectable fraction of PSR-BH binaries for three characteristic distances from earth namely 1 kpc, 5 kpc and 10 kpc in the top, middle and bottom respectively. Here, we see that a distance of 1 kpc, a large fraction of pulsars are detectable using higher-order searches like the template-bank pipeline. The median detectable fraction from the GPU accel.search pipeline is 30.2 %, the segmented accel. search pipeline is 44.7 % which increases to 72.5% using the template-bank pipeline. Large improvements in detectable fraction can be made here by improving the binary search algorithm. We call this regime algorithm-limited as compared to sensitivity limited for the 6 kpc case. In the case of 5 kpc, as shown in figure 3.24, the median detectable fraction is 2.5 %, 3.2% and 4.1% respectively whereas in the 10 kpc case it is 1.2, 1.4 and 1.8 % respectively. Based on our null detections, and our sensitivity limit tests, we think it is highly unlikely that a PSR-BH binary exists near the Galactic-plane at distances $d \leq 1$kpc. While our analysis, were restricted to orbital periods in the 4-24 hrs regime, orbits longer than a day should in principle be easier to detect due to their low line of sight acceleration value. A high eccentricity of such orbits could make them more difficult to detect close to periastron, however Bagchi et al. (2013) did report that on average highly

---

[34] Defined as the fraction of the sky covered by the pulsar radiation beam

[35] `https://www.skatelescope.org/`



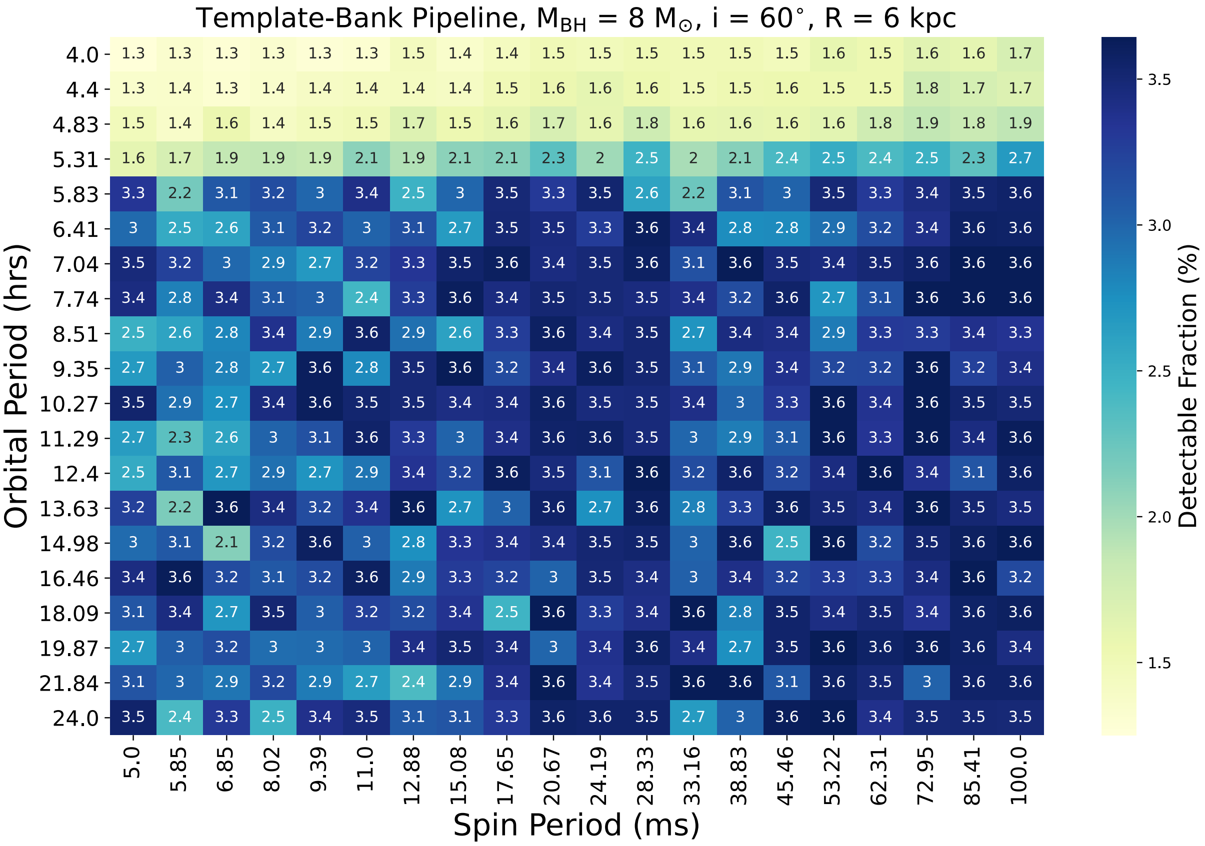
	5.0	5.85	6.85	8.02	9.39	11.0	12.88	15.08	17.65	20.67	24.19	28.33	33.16	38.83	45.46	53.22	62.31	72.95	85.41	100.0
4.0	0.69	0.58	0.64	0.6	0.66	0.62	0.51	0.57	0.51	0.47	0.48	0.46	0.47	0.47	0.49	0.42	0.44	0.4	0.41	0.36
4.4	0.58	0.57	0.62	0.58	0.54	0.52	0.52	0.53	0.47	0.43	0.41	0.44	0.47	0.45	0.44	0.45	0.44	0.35	0.36	0.38
4.83	0.48	0.56	0.43	0.53	0.49	0.47	0.39	0.45	0.41	0.37	0.41	0.34	0.42	0.41	0.42	0.4	0.35	0.32	0.33	0.32
5.31	0.41	0.37	0.32	0.32	0.33	0.26	0.3	0.26	0.26	0.22	0.28	0.19	0.29	0.26	0.2	0.19	0.21	0.19	0.22	0.17
5.83	0.12	0.24	0.13	0.13	0.14	0.11	0.19	0.14	0.11	0.12	0.11	0.18	0.23	0.13	0.14	0.11	0.12	0.11	0.11	0.1
6.41	0.14	0.19	0.18	0.13	0.12	0.14	0.13	0.17	0.11	0.11	0.12	0.1	0.11	0.16	0.16	0.14	0.13	0.11	0.1	0.1
7.04	0.11	0.12	0.14	0.15	0.17	0.12	0.12	0.1	0.1	0.11	0.11	0.1	0.13	0.1	0.11	0.11	0.11	0.1	0.1	0.1
7.74	0.11	0.15	0.11	0.13	0.14	0.2	0.12	0.1	0.11	0.11	0.11	0.11	0.11	0.13	0.1	0.17	0.13	0.1	0.1	0.1
8.51	0.19	0.18	0.15	0.11	0.15	0.1	0.14	0.17	0.12	0.1	0.11	0.11	0.16	0.11	0.11	0.15	0.12	0.12	0.11	0.12
9.35	0.17	0.14	0.15	0.17	0.1	0.16	0.11	0.1	0.13	0.11	0.1	0.11	0.13	0.15	0.11	0.12	0.12	0.1	0.12	0.11
10.27	0.11	0.14	0.16	0.11	0.1	0.11	0.11	0.11	0.11	0.1	0.11	0.11	0.11	0.14	0.12	0.1	0.12	0.1	0.11	0.11
11.29	0.17	0.22	0.18	0.14	0.13	0.1	0.12	0.14	0.11	0.1	0.1	0.11	0.14	0.15	0.13	0.1	0.12	0.1	0.11	0.1
12.4	0.19	0.13	0.17	0.14	0.16	0.15	0.12	0.12	0.1	0.11	0.13	0.1	0.13	0.1	0.12	0.11	0.1	0.11	0.13	0.1
13.63	0.12	0.25	0.1	0.11	0.13	0.11	0.1	0.16	0.14	0.1	0.16	0.1	0.15	0.12	0.1	0.11	0.11	0.1	0.11	0.11
14.98	0.14	0.13	0.26	0.13	0.1	0.14	0.16	0.12	0.11	0.11	0.11	0.11	0.14	0.1	0.2	0.1	0.13	0.11	0.1	0.1
16.46	0.11	0.1	0.13	0.13	0.13	0.1	0.15	0.12	0.12	0.14	0.11	0.11	0.14	0.11	0.12	0.12	0.12	0.11	0.1	0.12
18.09	0.13	0.11	0.17	0.11	0.14	0.12	0.13	0.12	0.2	0.1	0.12	0.11	0.1	0.15	0.1	0.11	0.11	0.12	0.1	0.1
19.87	0.17	0.13	0.13	0.14	0.14	0.14	0.11	0.11	0.11	0.14	0.11	0.1	0.11	0.16	0.1	0.1	0.1	0.1	0.1	0.11
21.84	0.13	0.14	0.15	0.12	0.15	0.17	0.21	0.14	0.11	0.1	0.12	0.11	0.1	0.1	0.13	0.1	0.11	0.14	0.1	0.1
24.0	0.1	0.2	0.12	0.19	0.11	0.11	0.13	0.13	0.12	0.1	0.1	0.11	0.17	0.14	0.1	0.1	0.11	0.11	0.11	0.11

Figure 3.22: Minimum Detectable Flux density towards PSR-BH binaries in the $P_{\rm orb}$-$P_{\rm spin}$ plane assuming a fixed black mass of 8 $M_\odot$. These limits were obtained from searches done using the template-bank pipeline.

eccentric orbits $e \sim 0.8$ are more easier to detect than mildly eccentric orbits $e \sim 0.1$ as the pulsar spends most of its time away from periastron. Due to the assumption of circular orbits in our searches, we cannot yet rule out the possibility of mildly eccentric orbits with a PSR orbiting a black hole with orbital periods $P_{\rm orb} >$ 4hrs or circular orbits with $P_{\rm orb} <$ 4hrs at nearby distances $d \leq$ 1kpc. However, since this regime is mostly limited by the search-algorithm used, future reprocessing efforts can certainly shed more light into the existence of nearby PSR-BH binaries and answer this question more definitively. For binaries located further away from us at distances at 5 and 10 kpc, our results show that algorithms can provide a marginal improvement in sensitivity, however a more significant gain can only be made by building more sensitive telescopes and improved electronics.

3.4 Concluding Remarks

Using the template-bank algorithm, we have conducted the most sensitive search for PSR-BH binaries in the southern hemisphere at L-Band using observations of the inner Galactic plane ($|b| < 3.5$) component of the HTRU-South survey. Our searches



	5.0	5.85	6.85	8.02	9.39	11.0	12.88	15.08	17.65	20.67	24.19	28.33	33.16	38.83	45.46	53.22	62.31	72.95	85.41	100.0
4.0	1.3	1.3	1.3	1.3	1.3	1.3	1.5	1.4	1.4	1.5	1.5	1.5	1.5	1.5	1.5	1.6	1.5	1.6	1.6	1.7
4.4	1.3	1.4	1.3	1.4	1.4	1.4	1.4	1.4	1.5	1.6	1.6	1.6	1.5	1.5	1.6	1.5	1.5	1.8	1.7	1.7
4.83	1.5	1.4	1.6	1.4	1.5	1.5	1.7	1.5	1.6	1.7	1.6	1.8	1.6	1.6	1.6	1.6	1.8	1.9	1.8	1.9
5.31	1.6	1.7	1.9	1.9	1.9	2.1	1.9	2.1	2.1	2.3	2	2.5	2	2.1	2.4	2.5	2.4	2.5	2.3	2.7
5.83	3.3	2.2	3.1	3.2	3	3.4	2.5	3	3.5	3.3	3.5	2.6	2.2	3.1	3	3.5	3.3	3.4	3.5	3.6
6.41	3	2.5	2.6	3.1	3.2	3	3.1	2.7	3.5	3.5	3.3	3.6	3.4	2.8	2.8	2.9	3.2	3.4	3.6	3.6
7.04	3.5	3.2	3	2.9	2.7	3.2	3.3	3.5	3.6	3.4	3.5	3.6	3.1	3.6	3.5	3.4	3.5	3.6	3.6	3.6
7.74	3.4	2.8	3.4	3.1	3	2.4	3.3	3.6	3.4	3.5	3.5	3.5	3.4	3.2	3.6	2.7	3.1	3.6	3.6	3.6
8.51	2.5	2.6	2.8	3.4	2.9	3.6	2.9	2.6	3.3	3.6	3.4	3.5	2.7	3.4	3.4	2.9	3.3	3.3	3.4	3.3
9.35	2.7	3	2.8	2.7	3.6	2.8	3.5	3.6	3.2	3.4	3.6	3.5	3.1	2.9	3.4	3.2	3.2	3.6	3.2	3.4
10.27	3.5	2.9	2.7	3.4	3.6	3.5	3.5	3.4	3.4	3.6	3.5	3.5	3.4	3	3.3	3.6	3.4	3.6	3.5	3.5
11.29	2.7	2.3	2.6	3	3.1	3.6	3.3	3	3.4	3.6	3.6	3.5	3	2.9	3.1	3.6	3.3	3.6	3.4	3.6
12.4	2.5	3.1	2.7	2.9	2.7	2.9	3.4	3.2	3.6	3.5	3.1	3.6	3.2	3.6	3.2	3.4	3.6	3.4	3.1	3.6
13.63	3.2	2.2	3.6	3.4	3.2	3.4	3.6	2.7	3	3.6	2.7	3.6	2.8	3.3	3.6	3.5	3.4	3.6	3.5	3.5
14.98	3	3.1	2.1	3.2	3.6	3	2.8	3.3	3.4	3.4	3.5	3.5	3	3.6	2.5	3.6	3.2	3.5	3.6	3.6
16.46	3.4	3.6	3.2	3.1	3.2	3.6	2.9	3.3	3.2	3	3.5	3.4	3	3.4	3.2	3.3	3.3	3.4	3.6	3.2
18.09	3.1	3.4	2.7	3.5	3	3.2	3.2	3.4	2.5	3.6	3.3	3.4	3.6	2.8	3.5	3.4	3.5	3.4	3.6	3.6
19.87	2.7	3	3.2	3	3	3	3.4	3.5	3.4	3	3.4	3.6	3.4	2.7	3.5	3.6	3.6	3.6	3.6	3.4
21.84	3.1	3	2.9	3.2	2.9	2.7	2.4	2.9	3.4	3.6	3.4	3.5	3.6	3.6	3.1	3.6	3.5	3	3.6	3.6
24.0	3.5	2.4	3.3	2.5	3.4	3.5	3.1	3.1	3.3	3.6	3.6	3.5	2.7	3	3.6	3.6	3.4	3.5	3.5	3.5

Figure 3.23: Detectable Fraction of PSR-BH binaries assuming a pseudo luminosity distribution (Faucher-Giguère & Kaspi, 2006). Here we assumed that these pulsars are located at a distance of 6 kpc from us which is the median NE2001 DM distance of all the new pulsar discoveries of LOWLAT.

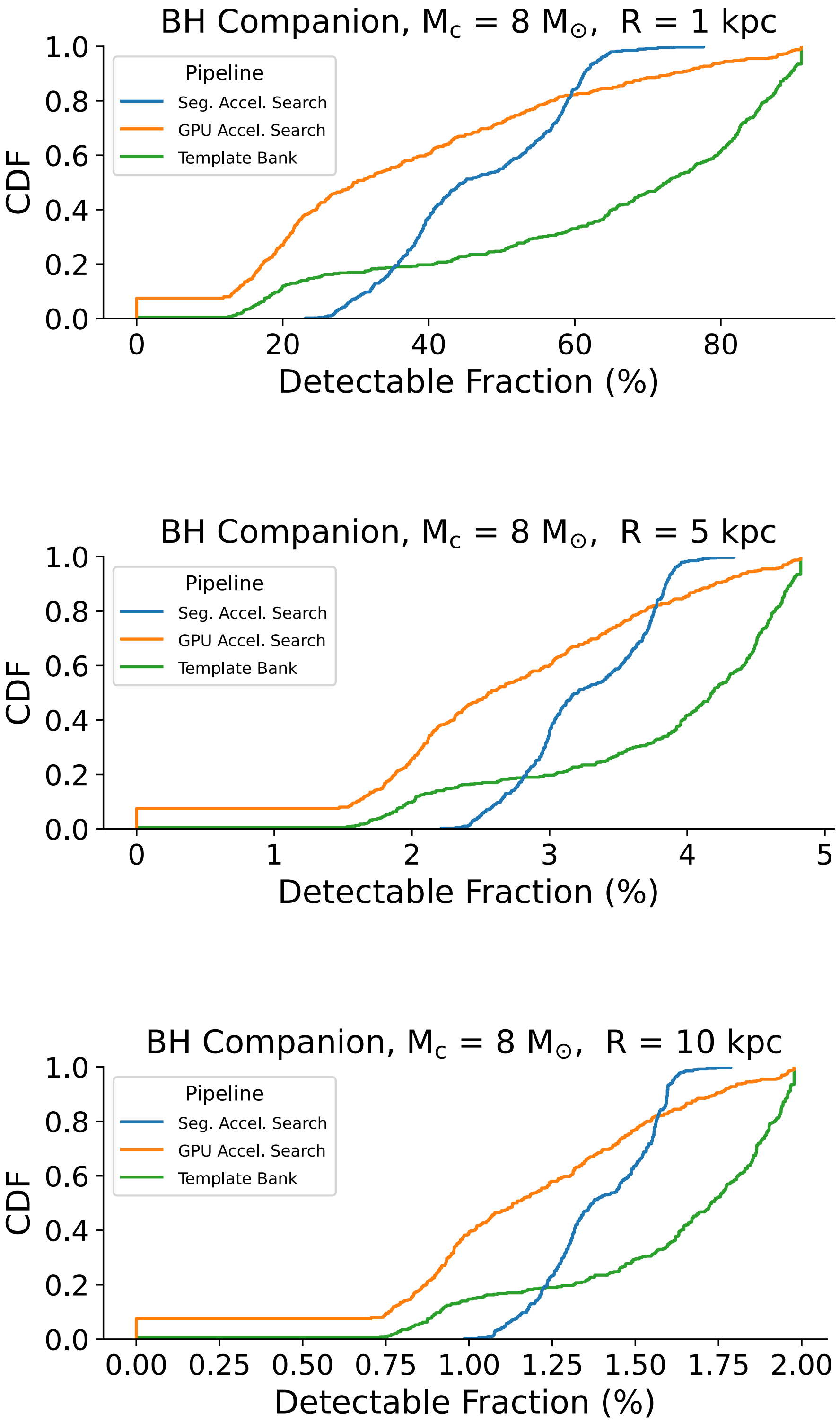


Figure 3.24: Cumulative Distribution Function (CDF) of detectable fraction of PSR-BH binaries located at different characteristic distances from earth. We show three cases here in the top, middle and bottom that correspond to 1 kpc, 5kpc and 10 kpc respectively. Additionally, we also show the detectability fraction from past-searches (blue and orange curve) in LOWLAT and compare them to searches done as a part of this thesis (green curve). See main text for discussion.

focused on the phase-space where 10-40% of the orbit is visible. Due to the increased amount of computational trials that was required to perform this search, a new GPU pipeline was developed as a part of this work, which was used to discover 70 new pulsars, 20 of which were initially found by our pipeline and the remaining were initially detected in other ongoing reprocessing efforts. Some of the discovery highlights include PSR J1743−24 which is a rare intermediate spin-period pulsar in a 70.7 day orbit around a light mass companion star. The formation history of this pulsar is likely inconsistent with standard evolutionary models and could be part of a separate population of pulsars with light companion masses and crucially exhibiting an intermediate spin-period. This pulsar also has a complicated pulse-profile and will require further polarisation-based investigation studies in order to better understand its emission geometry. PSR J1609−5225 has the highest detected DM in our discovery sample of $993(10)\,\mathrm{pc\,cm^{-3}}$. The integrated pulse profile of this pulsar shows a strong scattering tail and by modelling the trailing edge component of the pulse profile we measure a scattering timescale $\tau = 77 \pm 25\,\mathrm{ms}$ which is consistent with the predictions based on the NE2001 and YMW16 electron density models. We expect this pulsar to be a useful addition to the population of pulsars that are typically used for ISM related studies. PSR J1142−62 is a normal pulsar suspected to be in a binary orbit and if confirmed would be part of a minority population ($\sim 2\%$) of normal pulsars which have a binary companion. Additionally, we also presented an updated timing solution combining data from JBO and Parkes of PSR J1753−2819 which is a pulsar similar to PSR J1743−24 in its unusual intermediate spin-period but is in a much shorter orbit of $\sim 9.3\,\mathrm{h}$. In light of the high number of discoveries from the reprocessing, we also presented an updated survey yield analysis of the LOWLAT survey and compared properties of the new discovery population in the survey to the ones previously discovered. We find that on average we detected pulsars at a statistically significant lower spectral and folded S/N. This was possible due to careful analysis of the data preprocessing scripts like the downsampling routine and the dedispersion step size as well as gains in computational power. We find that the new discovery population in LOWAT presented in this thesis is on average at a higher DM than the ones previously reported in Ng et al. (2015); Cameron et al. (2020) and therefore its likely that we have probed a phase-space of pulsars located further away from us. Finally, we quantified the sensitivity of all the binary pulsar search pipelines used so far in LOWLAT towards PSR-BH binaries. We find that on average, our searches have atleast doubled our sensitivity towards PSR-BH binaries in the 6-12 hour orbital period range that was only possible due to improvements in the search algorithm presented here. Despite this, so far we report no detections for a PSR-BH binary or any new relativistic binary pulsar. We have used our non-detections to place constrains on the detectable fraction of PSR-BH binaries at 1, 5, 6 and 10 kpc respectively. Our results show that its highly unlikely that a PSR-BH binary exists near the Galactic-plane at distances $d \leq 1$kpc. in the $4-24\,\mathrm{h}$ orbital period range with a circular orbit. We cannot yet rule out the possibility of nearby PSR-BH binaries with mildly eccentric orbits or circular orbits shorter than 4 hours. Since, these regions are mostly limited by the algorithm and not sensitivity, future reprocessing efforts can shed more light into this. At distances further away from us $d \geq 5$kpc, we are largely sensitivity limited. While we can in principle double our detectability fractions towards PSR-BH binaries compared to acceleration searches, in terms of absolute numbers we are probing a relatively minor fraction of the total population. These regimes can only be fully probed by future pulsar surveys using more sensitive telescopes like MEERKAT and SKA. Some takeaways from our analysis for future processing strategies are to carefully check data preprocessing scripts, not ignore pulsar candidates at low FFT S/N and to increase the number of candidates that should be folded from any given beam. The computational feasibility of this strategy will largely depend on the search-setup as folding time will likely be the major bottleneck for future pulsar surveys. Alternatively, more work needs to be done to develop better RFI mitigation algorithms to reject spurious candidates. In terms of binary pulsar searches, a rich array of pulsar searches can still be done thanks to the development of our GPU pipeline which can be used for processing data from any radio pulsar survey. Some examples besides the one presented here include searches focused on finding PSR-WD binaries or DNS systems which will enable us to reduce our spin-period limit in the template-bank and help us to find more compact orbit binaries that are not detectable using an acceleration or jerk search.

CHAPTER 4

Pulsar Candidate Identification Using Semi-Supervised Generative Adversarial Networks.

This chapter is based on an article published in the Monthly Notices of the Royal Astronomical Society (Balakrishnan et al., 2021). I am the lead author of this publication. I wrote the machine learning pipeline presented here called SGAN and performed all the analysis including writing the article. The full list of the authors are given below:

Vishnu Balakrishnan,[1], David Champion,[1] Ewan Barr,[1] Michael Kramer,[1] Rahul Sengar,[2] and Matthew Bailes [2]

[1] *Max-Planck-Institut fur Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany*

[2] *Centre for Astrophysics and Supercomputing, Swinburne University of Technology, P.O. Box 218, Hawthorn, VIC 3122, Australia*

Abstract

Machine learning methods are increasingly helping astronomers identify new radio pulsars. However, they require a large amount of labelled data, which is time consuming to produce and biased. Here we describe a Semi-Supervised Generative Adversarial Network (SGAN) which achieves better classification performance than the standard supervised algorithms using majority unlabelled datasets. We achieved an accuracy and mean F-Score of 94.9% trained on only 100 labelled candidates and 5000 unlabelled candidates compared to our standard supervised baseline which scored at 81.1% and 82.7% respectively. Our final model trained on a much larger labelled dataset achieved an accuracy and mean F-score value of 99.2% and a recall rate of 99.7 %. This technique allows for high quality classification during the early stages of pulsar surveys on new instruments when limited labelled data is available. We open-source our work along with a new pulsar-candidate dataset produced from the High Time Resolution Universe - South Low Latitude Survey. This dataset has the largest number of pulsar detections of any public dataset and we hope it will be a valuable tool for benchmarking future machine learning models.

4.1 Introduction

Discovering a new pulsar can often lead to new and exciting science. Some examples include the discovery of PSR B1257+12 (Wolszczan & Frail, 1992), which led to the discovery of the first set of extrasolar planets. The first binary pulsar PSR B1913+16 (Hulse & Taylor, 1975b) and the subsequent measurement of its orbital period decay provided the first indirect evidence of gravitational waves. The discovery of the first pulsar triple system (a pulsar orbiting two white-dwarfs) led to one of the most stringent test of the Strong Equivalence Principle (SEP), a prediction of general relativity (Voisin et al., 2020). More recently, PSR J1141-6545 was used to infer Lense-Thirring precession (relativistic frame-dragging), a prediction of General Relativity (Venkatraman Krishnan et al., 2020). These examples are only some of the highlights that display the value of pulsar discoveries. Therefore, in order to keep pushing the boundaries of fundamental physics, it is important that we continue the investigation of new techniques in order to enhance the discovery process.

Identifying radio pulsars involves finding usually broadband periodic signals in noise-dominated data. As pulsar signals pass through the interstellar medium (ISM) before arriving at radio telescopes, their radio emission is “dispersed” by the free electron content in the ISM. The amount of dispersion is proportional to the dispersion measure (DM) which is proportional to the integrated column density of free electrons between the pulsar and the observer. This creates a frequency-dependant delay such that lower frequency signals arrive later compared to higher frequency signals. Since the true DM of a pulsar is a *priori* unknown, we typically de-disperse the data into multiple trial values. Once the data is dedispersed, periodic signals are identified in the timeseries by calculating a Fast Fourier Transform (FFT) (Cooley & Tukey, 1965) or by using the Fast Folding Algorithm (FFA) (Staelin, 1969). Once the top candidates have been identified, these signals are folded at their respective spin-period and dispersion measure to form a pulsar candidate. Pulsar candidates are four-dimensional data-cubes consisting of time, frequency, rotational phase and power of a signal. These are the end products that are produced by most FFT based pulsar-search pipelines[1]. The final step is usually performed manually. Pulsar candidates are visualized as series of diagnostic plots (see Section 4.2.6 for more details) which are used by pulsar astronomers to identify if a signal is from a genuine pulsar or not.

Modern pulsar surveys, like the High Time Resolution Universe Low Latitude survey (HTRU-S Lowlat) Keith et al. (2010) typically produce around 40 million pulsar candidates in one processing run. The increasing number of pulsar candidates can be attributed to multiple factors. Modern surveys tend to have higher time and frequency resolution. In addition, HTRU-S Lowlat also has a relatively long 72-minute integration time, which leads to larger FFTs and the requirement of additional acceleration/template-bank trails in order to be sensitive to binary pulsars. We refer the readers to Lyon et al. (2016) for a more in-depth review on this topic. Out of these 40 million candidates, only a few hundred thousand of them are expected to be

---

[1] Presto: `https://github.com/scottransom/presto`
Sigproc: `http://sigproc.sourceforge.net/`

real pulsar detections (multiple detections of known pulsars + new discoveries). This is because even after refining candidate lists to eliminate multiple occurrences of the same pulsar across many DM and acceleration/template bank trials and harmonics, many bright pulsars can appear in the sidelobes of many pointings and in a survey like HTRU-S Lowlat this results in a few hundred thousand detections of the known pulsars in the survey. Assuming an extremely optimistic average inspection time of one second per candidate, working 12 hours a day, it would take a human 2.5 years to go through the entire dataset. Future pulsar surveys using the Square Kilometre Array (SKA) telescope[2], are expected to increase this number further. Therefore, automated selection techniques that are optimised based on both speed and accuracy are of high importance for current and future pulsar surveys. Several papers have been written to address this topic. Eatough et al. (2010b) used twelve hand-crafted numerical features/scores to describe each pulsar candidate. These twelve features were then attached to a multi-layer perceptron to identify pulsars in the Parkes multi-beam pulsar survey (PMPS, Manchester et al. 2001). Lee et al. (2013) introduced a candidate ranking scheme based on six quality factors that were selected based on domain knowledge. Zhu et al. (2014) developed a Pulsar Image Classification System (PICS) which is an ensemble machine learning model based on Convolutional Neural Networks (CNN), Support Vector Machines (SVM) and Artificial Neural Network (ANN). This technique was trained on candidates from the Pulsar Arecibo L-band Feed Array (PALFA) survey (Cordes et al., 2006), and were successfully applied for identifying pulsars in the Green Bank North Celestial Cap (GBNCC) survey (Stovall et al., 2014). More recently, Guo et al. (2019), used a combination of deep convolution generative adversarial network and support vector machines (DCGAN + L2SVM) to achieve excellent results for candidates in the HTRU Medlat and PMPS survey. However, all these techniques require a large number of labelled pulsar candidates in order to perform well. In practice, since the number of pulsar detections is only a small fraction of the total candidates, (< 1 per cent), previous works either under-sample the number of non-pulsars in their training data or over-sample the pulsar detections. In this paper, we present results from training a machine learning algorithm to address the practical scenario where we typically have a small amount of labelled data along with a large amount of unlabelled data. This is called Semi-Supervised learning. Past applications using a similar approach in astronomy include applying Semi-Supervised learning on data from the Very Long Baseline Array (VLBA) Fast Radio Transients Experiment (V-FASTR) for radio pulsar candidate classification (Jones et al., 2012; Bue et al., 2014), applying a Semi-Supervised distributed algorithm called Co-Training, Distributed, Random Incremental Forest (CoDRIFt)[3] for single pulse pulsar candidate classification (Devine, 2020) and another study which uses a Semi-Supervised Deep Convolutional Neural network to classify radio galaxy images (Ma et al., 2019). We also compare our results to the purely supervised approach as done by previous works.

---

[2] `https://www.skatelescope.org/`

[3] `https://github.com/tdevine1/cluster/tree/master/code/CoDRIFt`

4.2 Methods

4.2.1 Machine Learning

Machine learning is a branch of computer science that deals with solving problems by learning through experience. In the classical setup, a human defines all the steps necessary for a computer to solve the problem. However, for complex tasks when it is not trivial to come up with a model to map the input data to our desired output, it is often desirable to learn from the data itself. This process of learning through experience is usually called "training" an algorithm.

There are broadly three classes of machine learning that are relevant for the work in this paper.

1. Supervised Learning: In supervised learning, we have data and its corresponding label, which in our case is a binary label between pulsar and non-pulsar signals. To the best of our knowledge, all the currently published papers in pulsar candidate classification fall under this category.

2. Semi-Supervised Learning: Semi-supervised learning is a branch of machine learning that combines a small amount of labelled data along with a large number of unlabelled data in order to obtain better learning performance. This is the problem we are trying to tackle in this paper.

3. Unsupervised Learning: In unsupervised learning, no labels are provided during training. It is up to the algorithm to find useful structure in the input data.

4.2.2 Artificial Neural Network (ANN)

ANNs are a class of supervised machine-learning algorithms that are commonly used for classification tasks. Variants of this network have been used previously in solving the pulsar candidate classification problem in Eatough et al. (2010b); Bates et al. (2012); Zhu et al. (2014); Bethapudi & Desai (2018). This algorithm has also been used in this paper for comparing our proposed architecture to the standard supervised learning case. We briefly summarise the different components of an ANN and its operation. For a more thorough explanation, refer to chapter five of Bishop (2006).

The simplest unit of an ANN is a neuron. These neurons are loosely inspired by biological neurons in the sense that there are input(s) to the neuron, an activation function and an output. Neurons are usually grouped together in layers. The first layer (often called the input layer) of the ANN is usually attached to the image or data we are interested to predict on and the last layer (often called output layer) is usually attached to the label we want to predict. Figure 4.1 is an example of single layer neural network, where there is one input layer, hidden layer and output layer respectively. Each neuron of a layer is connected to all the neurons of the next layer. A neural network with several hidden layers is usually referred to as a deep neural network or multi-layer perceptron (MLP). Assume we have an input vector $\vec{X}$ of 8 elements $\{X_1, X_2, ...X_8\}$ which is passed onto a neuron in the next layer. This neuron then calculates a weighted sum of all the values of $\vec{X}$ and applies a non-linear activation function on it. Mathematically, this output $\hat{y}$ can be written as:

$$\hat{y} = g(w_0 + \Sigma_{i=1}^{m} X_i w_i), \text{where,} \tag{4.1}$$

w_i is the weight of the i^{th} neuron which decides the relative importance of a neuron, w_0 is the bias term which is a trainable constant value for each layer, g is the activation function used. The purpose of an activation function is to decide if the neuron should be activated or not. This function helps in normalising the weighted sum values and additionally they introduce non-linearities to the network. Some of the activation functions used in this paper include the sigmoid function $\left(f(x) = \frac{1}{1+e^{-x}}\right)$, tanh $\left(f(x) = \frac{e^{-2x}-1}{e^{-2x}+1}\right)$ and Rectified Linear Unit 'ReLU' $(f(x) = max(0, x))$.

The process by which a neural network "learns" is by minimising a loss function. Loss functions are defined as the difference between the predicted output of a neural network to the ground truth labels. Since, we are dealing with a binary classification problem, the output of our neural network is a probability value between 0 and 1 for a candidate to be a pulsar. A value closer to either extremum indicates high confidence in our prediction. Our goal is to minimise the cross-entropy loss between our predicted labels and true labels. We use the standard softmax function $\left(\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}}\right)$ for converting values into probabilities. Here, $\vec{z}$ is the input vector passed on-to the softmax function and K is the number of classes in the classifer. In practice, these loss functions are minimised in an iterative fashion by calculating their negative gradients and propagating it backwards to the network using a process called back-propagation (Rumelhart et al., 1986).

4.2.3 Convolutional Neural Network (CNN)

CNNs are a class of machine learning algorithms that are commonly applied in the field of image classification. One of the earliest applications of CNNs was the LeNet-5 network, which was successfully used to recognise hand-written digits (LeCun et al., 1998). CNNs have also been successfully applied to the pulsar candidate classification problem previously in Zhu et al. (2014); Guo et al. (2019). An example of a CNN is given in Figure 4.2. The major difference here is that the fully connected layers have been replaced with convolutional layers. These convolutional filters act as feature extractors to identify important parts of the input image. The convolutional layer is typically followed by a max-pooling operation, where we find the maximum value of preceding neuron cluster and store them to a single neuron in the current layer. For example, if we have a 48x48x1 tensor, after a 2D max pooling operation of 2x2, the tensor's size changes to 24x24x1. This is done to constrain the dimensionality of the network while propagating only important information to the next layer. This is usually followed by a activation function, typically ReLU. Many such convolution, max-

---

[4] `https://alexlenail.me/NN-SVG/index.html`

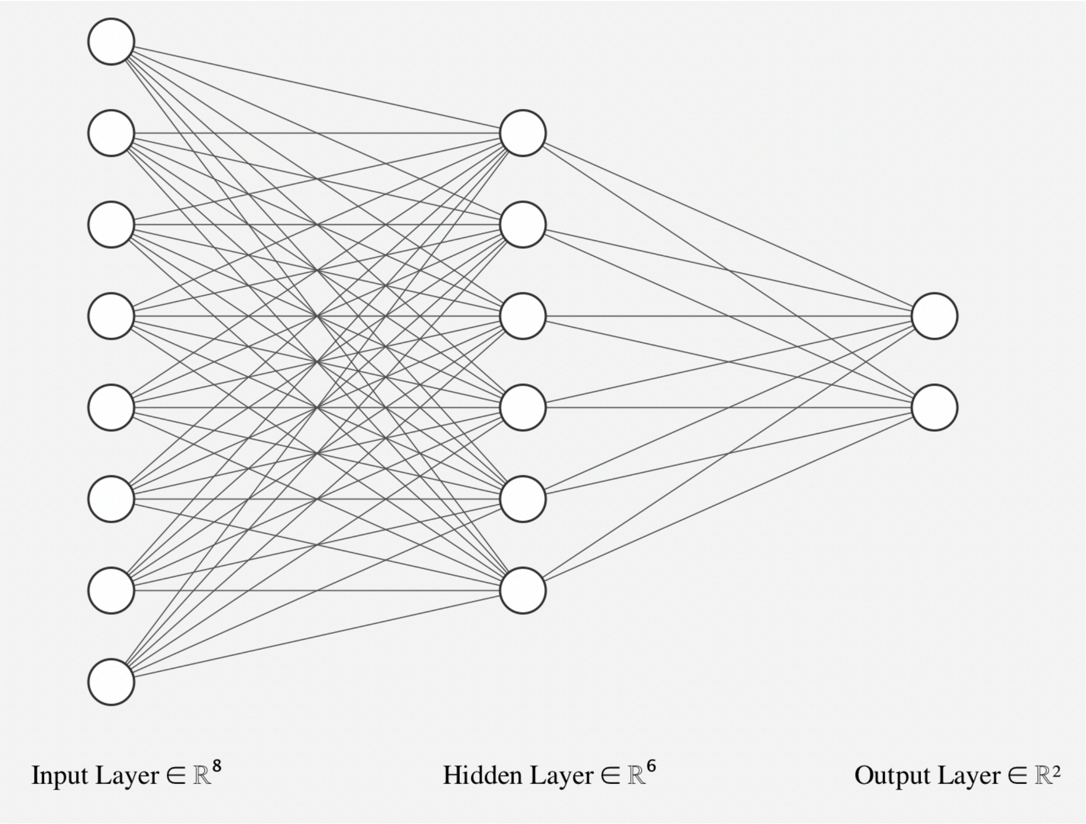


Figure 4.1: Multi-layer Artificial Neural Network with 8 input units, 6 hidden units and 2 output units. In shorthand, they are usually written as a 8:6:2 network. This diagram was created using an open source tool called NN-SVG[4](LeNail, 2019).

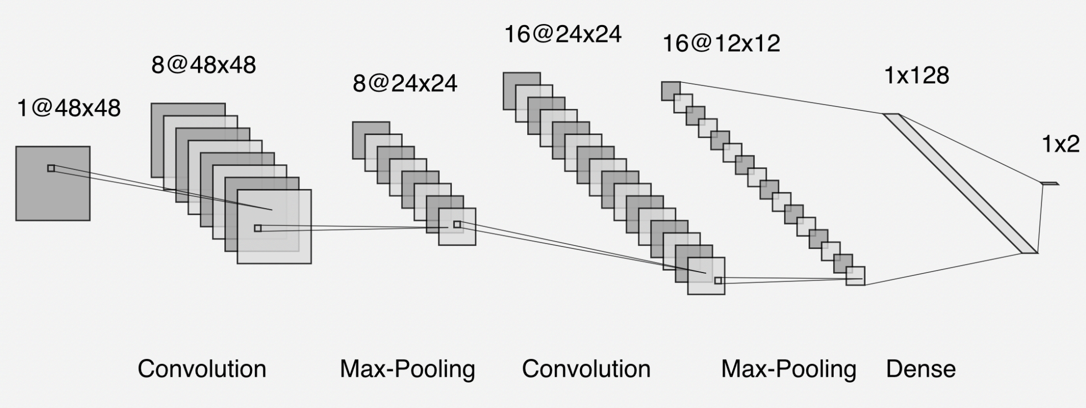


Figure 4.2: Schematic diagram of a typical CNN. Here the input image is a grey scale image of size 48x48 and the output contains two nodes for the binary class labels. 8 and 16 filters of size 4x4 were used for the convolutional layers. These act as feature extractors that identify import patterns in the input image. This diagram was created using an open source tool called NN-SVG [5](LeNail, 2019).

pooling and activation function layers can be concatenated together to a form a deep CNN. This is usually followed by a fully connected layer also known as a dense layer which is then finally connected to the output layer. The output layer for a classification problem is the amount of class labels we have in our data. Additionally for deep CNNs, a dropout layer is typically added after the max-pooling operation, which randomly drops off a certain percentage of the preceding nodes. This is done so that the neural network learns to better generalize its performance across the entire data. This is used as a regularization technique to avoid overfitting. CNNs are also trained using back-propagation. However, in practice since it is computationally difficult to calculate the gradient of the loss function for all images in the training data, we typically divide the data into mini-batches and use the stochastic gradient descent algorithm.

4.2.4 Generative adversarial network (GAN)

GANs are a class of machine learning algorithms (Goodfellow et al., 2014) in which two neural networks are trained simultaneously with opposing goals. They act against each other as adversaries in a minmax two-player game. A generative model G is tasked with generating new data that captures the distribution of the input data. A discriminator model D is tasked with classifying samples as either REAL (that belong to the original data distribution) or FAKE (samples generated by G). Assuming $V(D, G)$ is the value/loss function we are interested in then the first term in Equation 4.2, is the expectation of the logarithm of D's predictions when an input is from the real data sample. D's goal is to maximise this term. The second term represents one minus the expectation of the logarithm of D's predictions when data generated by G is passed onto to D. The goal of D is to maximise the second term but the goal of G is to

---

[5] See footnote 4.

minimise this term. This sets up the adversarial framework. In the ideal case, the generator perfectly samples the distribution of the input data distribution, and the discriminator output equals to 1/2.

$$\begin{aligned} \underset{G}{min}\underset{D}{max}V(D,G) = \mathbb{E}_{x\sim p_{data(x)}}[logD\ (x)] \\ +\mathbb{E}_{z\sim p_z(z)}[log(1-D(G(z)))] \end{aligned} \tag{4.2}$$

Algorithm 1: Minibatch stochastic gradient descent algorithm used for training a GAN.

```
for N training epochs do
    for k batches do
        Fix the weights of $G$ and update $D$
        Sample a mini-batch of m noise samples $\{z^{(1)},...,z^m\}$ from a noise
          distribution p$_z(z)$
        Sample a mini-batch of m real data samples $\{x^{(1)},...,x^m\}$ from the
          data distribution p$_{data(x)}$
        Update the Discriminator by ascending its stochastic gradient.
            $\nabla_{\theta_d}\frac{1}{m}\Sigma_{i=1}^{m}\left[logD(x^{(i)}+log(1-D(G(z^{(i)})))\right].$
          Fix the weights of $D$ and update $G$
        Sample a mini-batch of m noise samples $\{z^{(1)},...,z^m\}$ from a noise
          distribution p$_z(z)$.
        Update the Generator by descending it's stochastic gradient.
            $\nabla_{\theta_g}\frac{1}{m}\Sigma_{i=1}^{m}\left[log(1-D(G(z^{(i)})))\right].$
    end
end
```

In practice however, we use a minibatch stochastic gradient descent method and train the generator and the discriminator alternatively. The algorithm and the proof for the convergence of the algorithm can be found in Goodfellow et al. (2014), however for the benefit of the reader we briefly summarise the training of this network in algorithm 1.

GANs have been used successfully in a wide range of tasks ranging from computer vision, where they have been used for generating photo-realistic images (Karras et al., 2018), converting text to images (Reed et al., 2016), used as feature extractors for un-

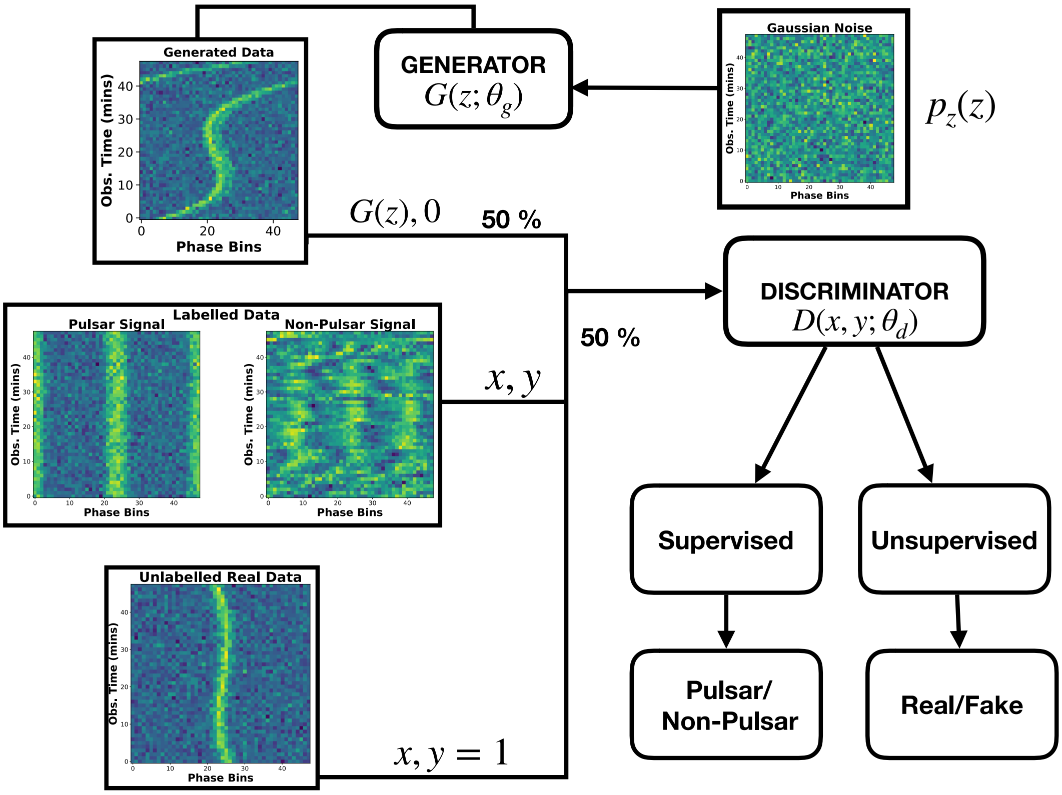


Figure 4.3: Schematic of the SGAN Architecture used in this paper. The generator is initialised with a noise (a.k.a latent vector) variable, which it then transforms to a fake generated image. The discriminator is fed images from three sources i) Labelled Pulsar candidates, ii) Unlabelled pulsar candidates provided with a positive label and iii) Fake generated images from the generator provided with a negative label. The discriminator is tasked with minimising both the supervised and unsupervised loss function.

supervised learning (Radford et al., 2015). In astronomy, GANs have been successfully demonstrated to recover features from astrophysical galaxy images beyond the deconvolution limit (Schawinski et al., 2017), create high-fidelity weak lensing convergence maps (Mustafa et al., 2019), modeling exoplanet atmospheres (Zingales & Waldmann, 2018) and more recently also in pulsar candidate identification (Guo et al., 2019). The standard framework of GAN described here usually comes under the category of unsupervised learning as no class labels are provided while training the network.

4.2.5 Semi-Supervised Generative Adversarial Network (SGAN)

SGANs are a variant of GAN where we can leverage both the ability of the generator to create realistic samples and the readily available unlabelled pulsar candidates to solve the semi-supervised classification problem. In the standard GAN problem, the output of D is a probability for the input image belonging to the training set. We modify this standard architecture slightly by adding the samples from G into our training set. We label them as a new generated class say K+1 where K is the total number of classes in our original classification problem. We then change the dimension of D's output from a binary classification output to a multi-class classification output {Pulsar, Non-Pulsar, Fake Data}. The main advantage of this technique is that we can now learn from our pulsar survey's unlabelled data.

There are three major components of this network, a supervised discriminator, an unsupervised discriminator and an unsupervised generator. The setup for the unsupervised discriminator and generator are similar to the standard GAN architecture discussed in Section 4.2.4. The supervised discriminator is provided with class labels (Pulsar or Non-Pulsar) that are available from our training set. The remaining unlabelled pulsar candidates were provided to the unsupervised discriminator with a positive label ('1') and generated fake candidates from G were provided with a negative label ('0'). For every training epoch, 50 per cent of the samples were taken from the generator and 50 per cent of the samples were taken from a combination of both labelled and unlabelled candidates. A schematic of this architecture can be found in Figure 4.3.

Mathematically, the loss function for SGANs can be written as

$$\begin{aligned}
\mathrm{L} &= -\mathbb{E}_{x,y\sim p_{data(x,y)}}[log\,\mathrm{p}_{\mathrm{model}}(y|x)] \\
&\quad -\mathbb{E}_{x\sim G}[\mathrm{log}\,\mathrm{p}_{\mathrm{model}}(y=K+1|x)] \\
&= \mathrm{L}_{\mathrm{supervised}} + \mathrm{L}_{\mathrm{unsupervised}}, \text{ where} \\
\mathrm{L}_{\mathrm{supervised}} &= -\mathbb{E}_{x,y\sim p_{data}(x,y)} log\,\mathrm{p}_{\mathrm{model}}(y|x, y<K+1) \\
\mathrm{L}_{\mathrm{unsupervised}} &= -\mathbb{E}_{x\sim p_{data}(x)} log\,[1-\mathrm{p}_{\mathrm{model}}(y=K+1|x)] \\
&\quad +\mathbb{E}_{x\sim G} log\,[\mathrm{p}_{\mathrm{model}}(y=K+1|x)].
\end{aligned} \tag{4.3}$$

The first term in the loss function is the so called supervised loss, $\mathrm{L}_{\mathrm{supervised}}$. This is similar to the standard loss function of any supervised classification model ($\mathrm{p}_{\mathrm{model}}$) where x is the data and y is it's corresponding label. The unsupervised loss term

($L_{\text{unsupervised}}$) consists of two parts, the first part corresponds to one minus the expectation that the model will output the new 'FAKE' class (K+1) given that the data is real and the second term is the expectation that the model will correctly identify the newly generated 'FAKE' class given that that the data comes from the generator. If we substitute, $D(x) = 1 - p_{\text{model}}(y = K + 1|x)$ into the unsupervised loss term, we notice that this is equivalent to the regular GAN's loss function described in Equation 4.2. The goal of SGANs is to minimise these two loss functions jointly. The key to the unsupervised term being useful is that the generator needs to be trained to approximate the input data distribution which in turn minimises the first term of the unsupervised loss function. The formalism for SGANs described here and several practical implementation tricks we used, were largely inspired by the work of Salimans et al. (2016).

4.2.6 Data Preprocessing and Features Used

A pulsar candidate is a four dimensional data cube of frequency channels, time, power and rotational phase of the signal. Since it is inconvenient to visualize four dimensional data, the convention is to plot various two-dimensional and one-dimensional projections of this data-cube to decide if a signal is really from a pulsar or not. The four feature plots pulsar astronomers most often use are:

1. Pulse Profile: This one dimensional intensity curve is created by integration over both time and frequency axes while preserving phase. Most real pulsars tend to have a one or multiple narrow peaks. However, there are some known exceptions. Some known pulsars, especially millisecond pulsars (MSPs) tend to have broader or close to sinusoidal profiles.

2. Frequency-phase Plot: This two dimensional plot is created by integrating over the time axis only. Real pulsars tend to be broadband, therefore, we expect a persistent bright signal (vertical line) across all sub-bands. However, pulsar scintillation caused by the interstellar medium can sometimes increase or decrease the signal in some frequency channels (e.g. PSR B0355+54 Xu et al. (2018)).

3. Time-Phase Plot: This two dimensional plot is created by integrating across the frequency axis only. We expect most pulsars to be persistent across observing time. There are some notable exceptions, for example a nulling pulsar like PSR J1727-2739 (Wen et al., 2016), relativistic binary pulsars which can have quadratic or cubic residuals in the time-phase plot or mildly accelerated pulsars where the acceleration falls between trial values.

4. DM-Curve: This is a one dimensional plot to find the best-fit dispersion measure value. In order to produce this, the candidate data is dedispersed around a few trial DM values from the DM used to fold the candidate. For each trial, it calculates the chi-squared of the dedispersed pulse profile against a horizontal line fit. A large chi-squared value is an indication that the signal deviates from white noise. Since pulsars are non-terrestrial signals, we expect the signal to peak at a

Table 4.1: Confusion Matrix for the Pulsar Candidate Classification problem.

True Label	Predicted Label: Pulsar	Predicted Label: Non-Pulsar
Pulsar	True Positive (TP)	False Positive (FP)
Non-Pulsar	False Negative (FN)	True Negative (TN)

non-zero DM value. The sharpness of the DM curve depends on the duty cycle of the pulsar.

We used the four features mentioned above to train the semi-supervised network. Before the data is passed onto the network, it is important to standardise the data, so that the algorithm is agnostic to spin-period, dispersion measure, observing frequency and integration time of an observation. We use the publicly available data pre-processing code made available by Zhu et al. (2014) for our work.[6] An example of the four features and the different types of signal in the training set is shown in figure 4.4. In order, to have the same number of bins for all candidates, the code downsamples and interpolates the data using linear interpolation for the 1-D plots and spline interpolation for the 2-D plots. The data is also normalized to have zero median and unit variance. We use 60 bins for the DM-curve and 64 bins for the pulse-profile. The time-phase and frequency-phase plots were resampled to a size of 48x48 bins. The bin sizes for the different features were chosen to maintain consistency with (Zhu et al., 2014) for our eventual comparisons.

4.2.7 Metrics Used

We use a combination of seven metrics to evaluate our different machine learning models. Since we have reformulated our network to a binary classification problem, there are four possible scenarios when we compare the predicted labels to the true labels. An example of this is shown in table 4.1. This is usually referred to as confusion matrix in the literature. Based on this we calculate the following metrics:

1. The simplest metric to calculate is accuracy. While this can be a useful metric to evaluate our model, care must be taken to ensure that our training data is balanced. In unbalanced training datasets, a high accuracy score alone is not an indication of a useful machine learning model.

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Predictions}} = \frac{\text{TP}+\text{TN}}{\text{TP}+\text{TN}+\text{FP}+\text{FN}}$$

2. Precision is defined as the fraction of true positives among all the positive label outputs of the model. This is a useful metric if our goal is to minimise the number of false positive (RFI + white noise) candidates which often translates

---

[6] https://github.com/zhuww/ubc_AI

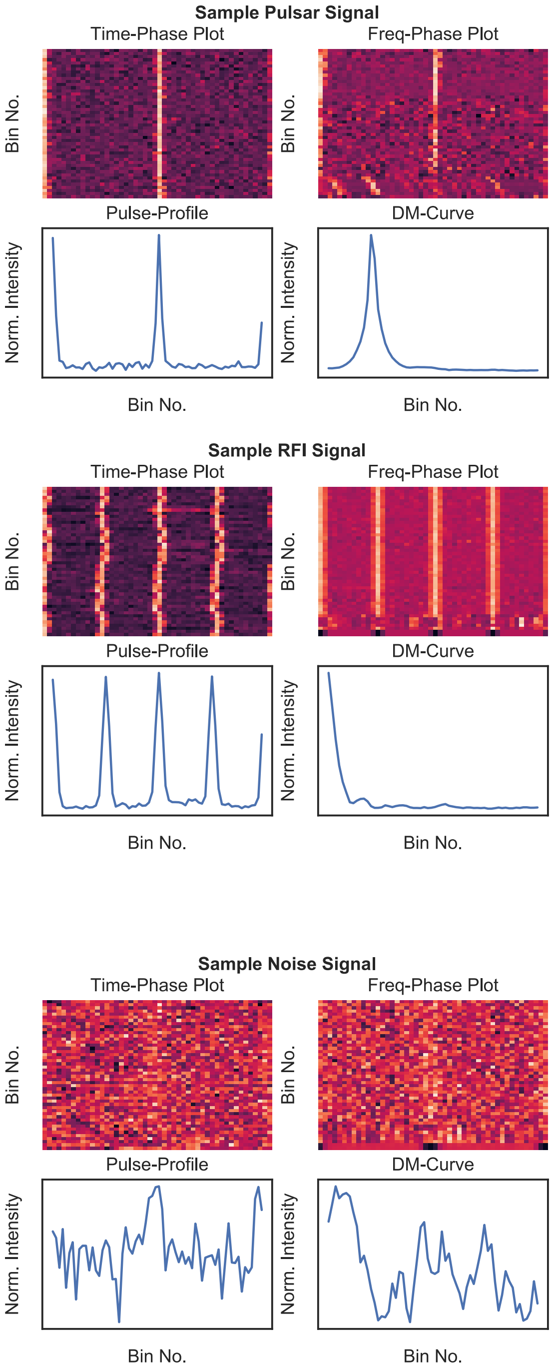


Figure 4.4: Sample data of the different types of signal in our training data: (a) Signal from a known pulsar, (b) Signal from a broadband RFI signal and (c) White Noise signal. Under each class label, we see a plot from the four features that was used to train all the machine learning algorithms used in this paper.

to less human hours spent in candidate inspection.

$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$

3. Recall or Sensitivity is the ratio of true positives to the total sum of true positives and false negatives. A high recall rate indicates that our model was successful in extracting most of the pulsar signals from our data.

   $\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$

4. F-score or F_1 score is defined as the harmonic mean of precision and recall.

   $\text{F-score} = 2\,\frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$

5. False positive rate (FPR) is analogous to precision. It is defined as the ratio of false positives to the total sum of false positives and true negatives. Unlike the other metrics used in this paper, a lower score of FPR is more desirable.

   $\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}$

6. Specificity is defined as the ratio of true negatives to the total sum of true negatives and false positives. This is analogous to the recall rate we defined earlier. A high specificity rate indicates that our model was successful in extracting most of the non-pulsar signals from our data.

   $\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$

7. G-Mean is defined as the geometric mean of recall and specificity.

   $\text{G-Mean} = \sqrt{2 \cdot \text{Recall} \cdot \text{Specificity}}$

4.3 Data Used in this study

We used observations from the High Time Resolution Universe South Low-Latitude Survey (HTRU-S Lowlat) to generate pulsar candidates. HTRU-S Lowlat is one part of the entire HTRU Survey that focuses on the inner galactic plane covering galactic longitude $-80^\circ < l < 30^\circ$ and galactic latitude $|b| < 3.5^\circ$. The observations were

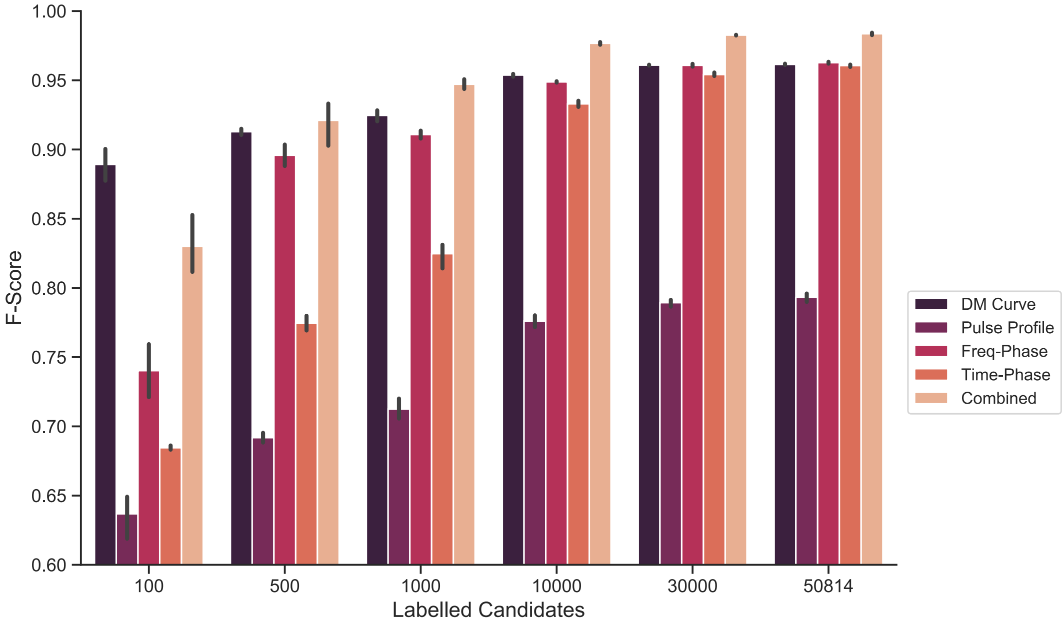


Figure 4.5: Mean F-score values of our supervised baseline model for varying number of labelled candidates. We find that "DM-Curve" is the best performing feature for labelled candidates below 500 and "Freq-Phase" as the best performing feature for larger number of labelled candidates. The combined model is a logistic regression model fit on the training data. Each model was averaged across five different batches of labelled data. The vertical bar displays the 95% confidence interval of our results.

Table 4.2: Total Number of pulsar and non-pulsar candidates in our labelled dataset. Candidates marked as 'Pulsar' also include harmonics and multiple detections of the same pulsar.

	Pulsar	Non-Pulsar	Total
Training Set	25,888	24,926	50,814
Validation Set	6,497	6,207	12,704
Testing Set	10,676	10,497	21,173

recorded for an integration time of 72 minutes with a frequency bandwidth of 400 MHz using the 64-m Parkes Radio Telescope. We refer the reader to Keith et al. (2010) for a full list of the observational setup and system configuration. To date, HTRU-S Lowlat has discovered >100 new pulsars. A full list of the initial discoveries and timing solution for these pulsars can be found in (Ng et al., 2015; Cameron et al., 2020). The pulsar candidates used for training all our models were generated from the re-processing of the HTRU-S Lowlat survey using the stochastic template-bank algorithm Harry et al. (2009), and folded using the PRESTO software suite Ransom (2011). The aim of the re-processing pipeline is to find compact relativistic binary pulsars that may have been missed by the first pass time-domain segmented acceleration search pipeline Ng et al. (2015). The total number of pulsar candidates produced by the re-processing pipeline for the entire survey is around 40 million. We selected 84,691 candidates that were labelled by eye to have approximately similar number of pulsar and non-pulsar candidates. We carefully chose pulsar candidates of different significance levels in order to create a diverse labelled candidate dataset. Our lowest detection significance of a true pulsar candidate is 4.3 sigma. The breakdown of candidates have been shown in Table 4.2. To the best of our knowledge, this labelled dataset has the largest number of pulsar detections out of all the publicly available pulsar candidate datasets. Labelled Pulsar candidates are critical for training machine learning algorithms as they are scarce ($<$ 1 per cent) compared to the total candidates produced in a pulsar survey.

4.4 Results

We start by splitting our entire labelled dataset into a train, validation and test dataset (60% train, 15% validation and 25% test, see table 4.2). The test dataset was never seen by the network while training. This dataset is only used in the end as a benchmark to evaluate all the different experimental setups described below. See figure 4.6, for the detection significance levels of pulsars in our test set. The validation dataset was used to tune hyperparameters of the different architectures and to select the best model during training. We train all the models separately on each of the four features described in section 4.2.6. Our software was built using Keras[7] (Chollet et al., 2015), a high-level open-source neural network library with Tensorflow 2.0 backend [8] (Abadi et al., 2015).

---

[7] `https://keras.io/`

[8] `https://www.tensorflow.org/`

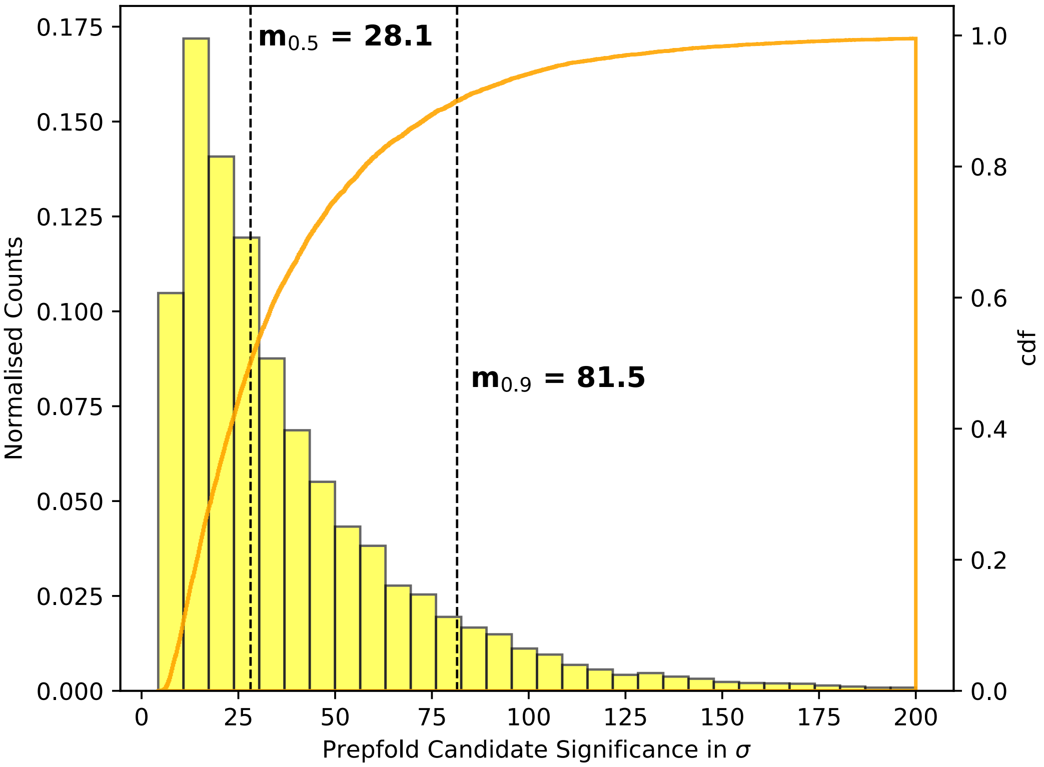


Figure 4.6: Histogram of the detection significance levels in sigma (i.e equivalent gaussian significance) of pulsars in our test dataset. The orange line shows the cumulative distribution function of the candidate significance levels. Vertical dashed lines indicate the median and 90th percentile values of the pulsar detection significance. For plotting purposes, we deliberately clipped out outlier candidates with a sigma greater than 200.

4.4.1 Supervised Learning Baseline

Our first goal was to build a model based on supervised learning that would act as our comparative baseline. For this, we trained a convolutional neural network (CNN) on the "Time-Phase" and "Freq-Phase" features and a multi-layer perceptron (MLP) for the "DM-Curve" and "Pulse Profile" features. Each network was trained for a total of 1000 epochs for different amounts of labelled data, saving only the model that produced the highest accuracy on our validation dataset. For each batch of labelled candidates, we split the training data to have an equal number of pulsar and non-pulsar signals. For example, 100 labelled candidates implies that the training data had 50 pulsar and 50 non-pulsar candidates. Since, the results are dependent on the subset of training samples used while training, we randomly selected five different combinations of labelled candidates and report the average values. The mean F-score performance of each of the four features is shown in Figure 4.5. We observe that in the regime of extremely limited labelled data (labelled candidates $<=$ 500), "DM Curve" acts as the best discriminator between pulsar and non-pulsar signals. However, as the number of labelled data increases, information about the persistence of the signal in "Freq-Phase" and "Time-Phase" regime become equally important. The individual score from each feature were combined using a Logistic regression model with "L2" regularization. This is marked as the combined model. Ideally, the combined model should be the best performing model. This holds true for our experiments with labelled candidates greater than five hundred. However, our combined model performs worse in the low labelled data regime (labelled candidates $=$ 100) because the models trained on "Pulse-Profile", "Time-Phase" and "Freq-Phase" brings down the net average performance of the network.

4.4.2 Semi-supervised GAN

4.4.2.1 Model Architecture and Implementation Details

There are three major components to an SGAN Network. A supervised discriminator, an unsupervised discriminator and an unsupervised generator. The simplest implementation is to have a single discriminator with multiple output layers. The first output layer solves the unsupervised task and outputs if the data is REAL/FAKE and the second layer solves the supervised task and outputs if the signal is from a pulsar or not. The drawback of this approach is that when we pass unlabelled candidates from the generator, there is no supervised label associated with them. Hence, this creates the need for an extra 'FAKE' class label for the supervised classifier. In this paper, we follow the technique described in Salimans et al. (2016) which removes the need for an extra class label. In this case, we built two separate models for the supervised and unsupervised task. Both models share the same feature extraction layers. However, the supervised model is attached to a softmax activation function whereas the unsupervised model takes the output from the supervised model prior to the activation function and calculates a normalized sum of exponential outputs. This custom activation function for the unsupervised discriminator $D(x)$ forces the model to give a strong prediction for real samples and lower values for the generated fake samples

$$D(x) = \frac{Z(x)}{Z(x)+1}, \text{where } Z(x) = \Sigma_{k=1}^{K} exp[l_k(x)]. \tag{4.4}$$

Our work is built on top of an open-source implementation of SGAN networks for MNIST digits[9]. We extensively modified the discriminator and generator architecture in order to get better results for our data. The discriminator architecture is similar to the CNN model used for the supervised baseline model. We obtained better results with larger convolutional kernels of size 7x7 compared to the 3x3 kernels that worked well for the supervised baseline models. The discriminator trained on the “DM Curve” and “Pulse-Profile” was a 1-D convolutional neural network with a convolutional kernel size of 7. Additionally, we used max-pooling to down-sample our images instead of stride convolutions. For the generator, we perform the transpose of convolutions in order to create images that are fed into the discriminator. Additionally, we also do batch normalization in order to speed up the training of the generator. The generator for the 1-D data was a multi-layer fully connected dense neural network. We use the tanh function as the activation function for the output layer of the generator. GANs can easily suffer from overconfidence. Therefore, as a regularization technique, we used soft and noisy labels while training. This means that if a candidate is real, instead of giving the label a value equals 1, we give a value in a range between 0.7-1.2 for the 2-D features and a value between 0.9-1 for the 1-D features. Around 5 per cent of the time, we intentionally flip the labels, we found that this helps to improve the overall performance. Our best performing model uses the Adam Optimizer (Kingma & Ba, 2014) with a learning rate of $\eta = 0.0002$ and $\beta_1 = 0.5$.

4.4.2.2 Effect of unlabelled data

In order to test if the SGAN network can learn from unlabelled data, we split the training data-set into smaller groups ranging from 100, to 30,000 candidates similar to our supervised learning baseline model experiments. Similarly, the number of unlabelled candidates used while training was also varied from 0 to 20,000. We trained the SGAN network for 400 epochs in each configuration, saving only the model that produced the best results on the validation dataset. Since, the results are dependent on the subset of training samples used while training, we randomly selected five different combinations of labelled and unlabelled datasets and report the average values. The mean F-score values of SGAN trained on the all four features is shown in Figure 4.7. We observe that increasing the number of labelled candidates in the training set drastically improves the final performance of the network. In addition, we clearly see that unlabelled candidates also improve the overall performance of the network. This effect is particularly significant in the low labelled data regime. For example with 100 labelled candidates in the training set, the unlabelled data improved the F-score of the network by at least 6 % for all features including an improvement of 12 % on the network trained on the

---

[9] https://machinelearningmastery.com/semi-supervised-generative-adversarial-network/

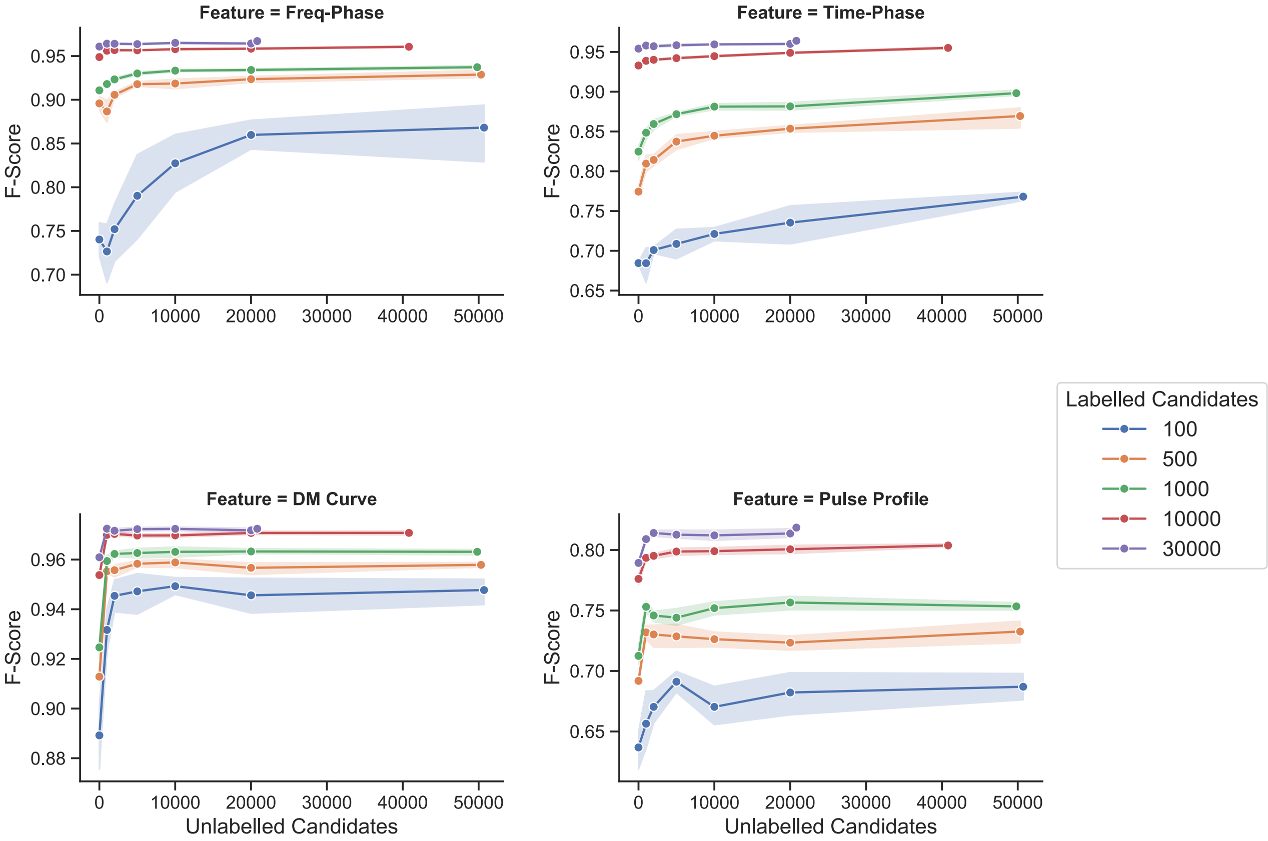


Figure 4.7: Mean F-score performance of each of the four features trained for 400 epochs with varying amounts of labelled and unlabelled samples. The interval around each line indicates the 95 % confidence interval associated to each value. This figure demonstrates that unlabelled candidates can be leveraged to improve the overall performance of a machine learning algorithm. This effect is particularly dominant with labelled candidates ≤ 1000. The boost in performance of the “Pulse-Profile” and “DM-Curve” feature for labelled candidates ≥ 10000, was critical to improve the overall performance of the combined model. Similar improvements were also seen for all metrics defined in section 4.2.7. These results can be found in table D.1.

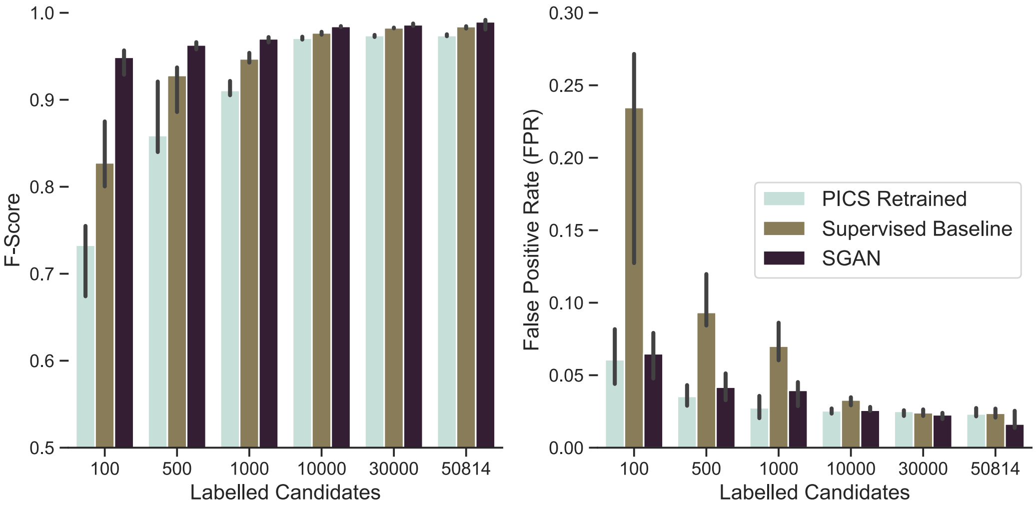


Figure 4.8: Mean F-score performance (left) and False-Positive Rate (right) of the Combined Semi-Supervised GAN (SGAN), our Supervised Baseline Model and the re-trained version of Pulsar Image Classification System (PICS). The black vertical line indicates the 95% confidence interval of our values. These results indicate that the SGAN architecture provides better classification results compared to both the supervised machine learning algorithms for all combinations of labelled data. For labelled candidates below 1000, the re-trained version of PICS has a lower false-positive rate compared to SGAN, however this comes at the cost of a significantly lower F-Score, thus making it a less desirable model. The difference in performance between the supervised baseline model and the retrained version of PICS can be partially attributed to the fact that PICS was not re-trained on thee same validation dataset. For labelled candidates below 1000, each of the supervised model ended up optimising for different metrics. Our supervised model resulted in a better recall rate and F-score at the cost of a worse false-positive rate and specificity score. We refer the readers to the appendix D.1 for the scores across all metrics. The SGAN model provides the overall maximum gain when there are fewer labelled candidates ($<$ 1000) available.

"Freq-Phase" feature. As the number of labelled candidates increase, we see that the semi-supervised classifier still provides better results. However, the performance boost provided by unlabelled candidates is significantly lower. We believe the reason for this is two-fold. With larger amount of labelled data, there is little room for improvement as the network has already learnt a good solution to solve the pulsar candidate identification problem. The second reason is that in order to fully utilise the strengths of the semi-supervised algorithm, we need to use a significantly large fraction of unlabelled candidates compared to the labelled candidates. Our final model which was trained on a much larger unlabelled candidate database has been described in Section 4.4.4.

4.4.3 Comparing SGAN with supervised models

In this section, we compare the performance of our two-layer ensemble SGAN network to the ensemble standard supervised baseline algorithm described earlier as well as a re-trained version of the Pulsar Image Classification system (PICS) (Zhu et al., 2014). In all cases, the individual scores from each of the four features were combined using a Logistic regression model with L2 regularization. We trained all three networks with the same labelled candidates for each experiment. In addition, unlabelled candidates were also used to trained the SGAN model. The same validation dataset was used to tune hyperparameters for the supervised model and the SGAN model. We didn't use a validation set for re-training PICS. This was because there wasn't a provision to provide a validation dataset in the re-training script provided by (Zhu et al., 2014). We presume that PICS trained on minimising the overall training loss. We find that the ensemble SGAN outperforms the standard supervised baseline algorithm as well as re-trained version of PICS for all combinations of labelled datasets and based on all the metrics discussed in Section 4.2.7 including higher accuracy, precision, recall rates and a lower false positive rate. For reasons of brevity we only show the mean F-score and False-Positive Rate (FPR) values in Figure 4.8. The full table comparing results of all the metrics can be found in Table D.1 in the appendix.

4.4.4 Best Performing Model

In this section, we describe our best performing model that was trained using the entire training set plus 265,172 unlabelled candidates. Results from five different training runs from the best performing semi-supervised and supervised models are summarised in Table 4.3. The confusion matrix of the predictions of this model on the test set is shown in Table 4.4. Our best model achieved an overall F-score of 99.2 %, recall rate of 99.7% and a false positive rate of 1.63%. Our best performing model has been merged into the HTRU-S Lowlat survey post-processing pipeline, and has already discovered eighteen new pulsars. These new pulsars had a detection significance ranging from 5.8 - 19 sigma. The SGAN network played a crucial role in discovering the lower detection significance pulsars as they are usually buried inside several non-pulsar candidates. A full list of these pulsars with their respective Spin-Period, DM and timing solutions will be the subject of a future publication.

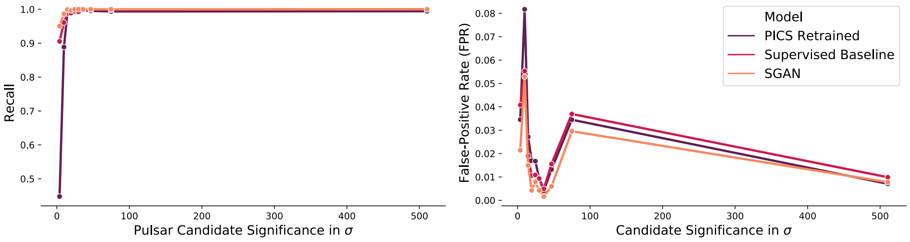


Figure 4.9: Recall rate (left) and False-Positive rate (FPR) (right) from the best-performing models of the three networks across different candidate detection significance levels (SNR). For the x-axis, we divided the candidates in our test dataset into ten quantile regions based on their detection significance level with a similar number of candidates across each bin. As expected, all three models perform better when the significance level of the candidate is higher. The re-trained version of PICS suffers from a large performance loss (recall: 0.45) at lower candidate significance levels (0-7.4 sigma). The SGAN network also suffers a small performance loss, however it still does better compared to the other models. The FPR for all three networks are much more interesting. As expected, we see a large FPR when the detection significance is lower, followed by a lower FPR rate. However, we again see an increase in the FPR rate at high significance levels. This is mostly caused by bright broadband RFI signals which look like pulsars and are detected with very high significance levels. The FPR rate is highest for significance levels between 7.4-12.1 sigma. These are mostly weak pulsar-like signals that are caused white-noise lining up to look like pulsars. See table D.2 for the performance of the network across other metrics.

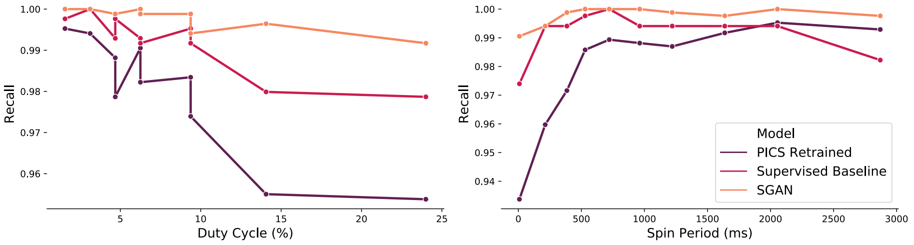


Figure 4.10: Recall rate of all three networks across different pulsar duty cycle ranges (left) and Spin-Period (right). The x-axis of these plots were made by dividing the duty cycle and spin-period of all pulsars in our test dataset across ten quantile regions such that they have a similar number of pulsars in each bin. It is easier to spot narrow duty cycle pulsars by eye, we see a similar effect with our trained neural networks. The magnitude of difference however varies across each model and is significantly lower for the SGAN model. Our models also tend to do slightly better for slow pulsars. However, both these effects could be correlated as pulsars with a slower-spin period also tend to have a narrow duty cycle.

Table 4.3: Results comparing the best performing combined supervised model trained on the entire training set against the best performing combined SGAN model trained on the entire training set plus 265,172 unlabelled candidates. Values reported in this table are the mean and standard deviation after repeating the training run five times.

	PICS Retrained	Supervised Baseline	SGAN
Accuracy	0.973 ± 0.001	0.984 ± 0.001	0.989 ± 0.004
Precision	0.977 ± 0.002	0.977 ± 0.002	0.984 ± 0.004
Recall	0.970 ± 0.001	0.991 ± 0.001	0.994 ± 0.004
F-score	0.973 ± 0.001	0.984 ± 0.001	0.989 ± 0.004
Specificity	0.976 ± 0.002	0.976 ± 0.002	0.983 ± 0.004
G-Mean	0.973 ± 0.001	0.983 ± 0.001	0.989 ± 0.004
FPR	0.023 ± 0.002	0.023 ± 0.002	0.016 ± 0.004

Table 4.4: Normalised confusion matrix of the predictions of the best ensemble SGAN model on the test set.

		Predicted	
		Pulsar	Non-Pulsar
True	Pulsar	0.997	0.003
	Non-Pulsar	0.014	0.986

4.4.5 Performance across Detection Significance, Duty cycle and Spin Period

In this section, we briefly analyse the performance of the best model from all three networks with various pulsar parameters like detection significance, duty cycle and spin-period. We start by splitting all the candidates in our test dataset based on their detection significance into ten quantile regions such that each bin has a similar number of candidates. We then calculate the performance of the best performing model from each of the three networks described earlier in this bin range. This is shown in Figure 4.9. As expected, we see an improvement in the recall rate of the network for higher detection significance. The performance drop for the re-trained version of PICS is drastic as it drops to a recall rate of 0.45 for pulsars with sigma values between 0 and 7.4. Both the supervised model and SGAN model do significantly better here scoring at 90.5 % and 95.0% respectively. The False Positive rate (FPR) is more interesting. As expected, we see a higher FPR at lower significance levels and visc versa. However, we see an increase in FPR rate for very high candidate significance levels (95.9 - 925.2 sigma). This was caused by bright broadband RFI signals that look very similar to pulsar signals. The FPR rate is highest for candidates with significance levels between 7.4 and 12.1 sigma. These are mostly caused by weak pulsar-like signals formed when white-noise lines up by chance. See table D.2 for the performance of the network across additional metrics.

We also explored the performance of the neural network across the duty cycle of the pulsars in our test dataset. Similar to the previous experiment, we divide the candidates in the test set based on their duty cycle into ten quantile regions such that there are similar number of candidates in each bin. For a human, it is easier to spot pulsars with a narrow duty cycle. We see a similar trend in the performance of the neural networks as well. However, the performance drop is not as drastic as compared to the detection significance level. We also repeated this experiment across different spin-period ranges. Both of these are shown in Figure 4.10. We see that pulsars with a slower spin-period appear to be slightly easier for the neural network to find. However, these two effects are highly correlated as slow spin-period pulsars also tend to have narrow duty cycle ranges.

4.5 Discussion

4.5.1 Inference Speed

Since our software has been built using Keras and Tensorflow, the final trained models can be used to evaluate pulsar candidates both on a CPU or on any Nvidia GPU. The inference rate of our model, benchmarked on a single Nvidia Tesla P100 GPU, is 5.22 ± 0.01 ms using a batch size of 20,000. This makes our model particularly suitable for implementing in a blind survey like the HTRU-S Lowlat as the entire batch of 40 million candidates can be scored in $\approx$ 58 hours on a single GPU. Additionally, our architecture can be easily retrained and redeployed as more labelled data become available. Our software can be found on GitHub[10]. We also provide a Dockerfile which can be used to create a Docker image in order to ensure easy reproducibility of our results.

4.5.2 Training Time and suitability for future pulsar surveys

The training time for SGAN is considerably longer compared to training a standard supervised deep learning architecture like a CNN or an ANN. For example, our supervised baseline pipeline on average took less than an hour to finish training for 1000 epochs on a Nvidia Tesla P100 GPU whereas our final model from the SGAN architecture took about four hours to train for 400 epochs. The main reason for this is because we are now training two neural networks alternatively and using a considerably larger sample of data (unlabelled candidates) while training. Keras and Tensorflow currently support training on multiple GPUs which can help reduce the net training time. While the training time is still acceptable for our needs, this architecture may not be the best approach for online data processing where the model needs to be re-trained in quasi real time. We refer the readers to the work of Lyon et al. (2016) that focuses more on speed rather than final classifier performance. The advantage of using our proposed architecture is higher performance because we can learn from unlabelled candidates. This is

---

[10] `https://github.com/vishnubk/sgan`

especially useful for RFI-rejection as such signals can have different signatures depending on the source (aircraft navigation, mobile phone, WiFi, satellites). Additionally, the RFI environment near a telescope is expected to change with the advancement of fifth generation wireless technology and therefore the ability to have a system that can adapt on relatively short timescales with high performance has huge value. Additionally, this technique also helps minimise the number of labelled candidates which saves human hours required to achieve satisfactory performance. The amount of unlabelled candidates that needs to be used while training depends on the number of labelled candidates available and is usually a trade-off between performance and training time. In our experiments, we achieved improved performance by having at least 5000 unlabelled candidates when the number of labelled candidates were below 1000. We observe that with 50,814 labels, we needed more than 200,000 unlabelled candidates to notice an improvement. It is also important to experiment with different amounts of unlabelled data as sometimes having more can make the model perform worse. We believe that using a static trained supervised model for classifying candidates from future pulsar surveys may not be the most optimal approach. Since labelling millions of candidates is not a scalable solution, we hope more attention goes into solving the pulsar candidate classification problem using a combination of labelled and unlabelled candidates.

4.5.3 Future Work

In this section, we briefly discuss some techniques that can be used to improve on our current models.

4.5.3.1 Improving the Supervised Baseline Model

The performance of our supervised baseline model in the regime of high labelled candidates can be improved using a much deeper convolutional neural network. Large Networks pretrained on ImageNet like VGG16 (Simonyan & Zisserman (2014)), InceptionV3 (Szegedy et al. (2015)) and ResNet50 (He et al. (2015)) among others can be used and their final few layers can be re-trained on a pulsar candidate dataset. This technique is called transfer learning and it has been successfully employed in various computer vision tasks including classifying Fast Radio Bursts (FRB) and RFI (Agarwal et al., 2019). In order to have a fair comparison between such networks and SGAN, we propose using a similar deep architecture for the discriminator of SGAN and comparing their performance.

4.5.3.2 Improving the SGAN Model

We believe that the performance of our SGAN model can be improved further by using a technique called feature matching. For this, we change the loss function of the generator such that its goal changes from beating the discriminator to minimizing the statistical difference between real and generated images. We refer the readers to Salimans et al. (2016) for a more detailed explanation of this technique. Another technique to improve the final semi-supervised classification accuracy is to use a Bad GAN. Instead of training towards a perfect generator, which produces images indistinguishable from real images, our goal in this architecture is to generate data that complements data produced from the discriminator. The drawback of this approach is that the quality of the generated images in general would be worse but this architecture has been shown to provide better classification results on the MNIST, SVHN and CIFAR-10 datasets (Dai et al., 2017).

4.6 Conclusion

In this paper we use an ensemble Semi-Supervised Generative Adversarial (SGAN) framework to classify pulsar candidates in the HTRU-S Lowlat Survey. We demonstrate that this algorithm achieves an overall F-score of 99.2% on our dataset and outperforms the standard supervised baseline algorithm and the re-trained version of PICS. The performance difference between both the techniques is significant in the low labelled-candidate regime. SGAN achieved a recall rate of 96.0% with 100 labelled-candidates compared to 85.6% from our supervised baseline model and 60.3 % from the retrained version of PICS. The main advantage of our proposed network is the ability to leverage readily available unlabelled candidates for achieving better results. We believe this technique will be even more useful for future pulsar surveys as the number of pulsar candidates scale up and maintaining a large labelled dataset becomes increasingly challenging. Our architectures are frequency and telescope agnostic, therefore they can be in principle applied to other ongoing pulsar surveys. We additionally share our code, and a dockerfile to enable reproduciblity of our work.

Acknowledgements

Observational data used in this work were made available by High Time Resolution Universe (HTRU) scientific collaboration. The Parkes Observatory, used in the collection of this data is part of the Australia Telescope National Facility which is funded by the Australian Government for operation as a National Facility managed by CSIRO. The data analysis were performed on the OzSTAR national supercomputing facilities at Swinburne University of Technology and the HERCULES computing cluster operated by the Max Planck Computing & Data Facility (MPCDF). OzSTAR is funded under Astronomy National Collaborative Research Infrastructure Strategy (NCRIS) Program via Astronomy Australia Ltd (AAL). We would like to thank members of the open-source community for maintaining packages that were directly used for our work including NumPy Oliphant (2006), Matplotlib Hunter (2007), Seaborn Waskom et al. (2017), Scikit-learn Pedregosa et al. (2011), Keras (Chollet et al., 2015) and Tensorflow (Abadi et al., 2015).

Chapter 5

Coherent Search for Binary Pulsars across all Five Keplerian Parameters in Radio Observations using the template-bank algorithm

This chapter is based on an article that has been published in the Monthly Notices of the Royal Astronomical Society (Balakrishnan et al., 2022). I am the lead author of this article. I developed the five-parameter search algorithm and pipeline presented here and performed all the analysis including writing the article. This chapter is based on the submitted version of the paper to MNRAS and hence is outdated. Please refer to the reference given above for the accepted version of the paper with minor changes. The full list of the authors are given below:

Vishnu Balakrishnan,[1], David Champion,[1] Ewan Barr,[1] Michael Kramer,[1],V. Venkatraman Krishnan,[1],Ralph P. Eatough,[1,2], Rahul Sengar,[3] and Matthew Bailes [3]

[1] *Max-Planck-Institut fur Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany*

[2] *National Astronomical Observatories, Chinese Academy of Sciences, 20A Datun Road, Chaoyang District, Beijing 100101, PR China*

[3] *Centre for Astrophysics and Supercomputing, Swinburne University of Technology, P.O. Box 218, Hawthorn, VIC 3122, Australia*

Abstract

Relativistic binary pulsars orbiting white-dwarfs and neutron stars have already provided excellent tests of gravity. However, despite observational efforts, a pulsar orbiting a black hole has remained elusive. One possible explanation is the extreme Doppler smearing caused by the pulsar's orbital motion which changes its apparent spin frequency during an observation. The classical solution to this problem has been to assume a constant acceleration or jerk for the entire observation. However, this assumption breaks down when the observation samples a large fraction of the orbit. This limits the length of search observations, and hence their sensitivity. This provides a strong motivation to develop techniques that can find more compact binaries in longer

observations. Building on work done by the gravitational wave search community, we present a GPU-based radio pulsar search pipeline that can perform a coherent search for binary pulsars by directly searching over three or five Keplerian parameters using the template-bank algorithm. We compare the sensitivity obtained from our pipeline with standard acceleration and jerk search pipelines for simulated pulsar-stellar-mass black hole binaries and observations of the double pulsar J0737-3039A. We also comment on the computational feasibility of our pipeline for blind pulsar surveys and targeted searches. Our benchmarks indicate that circular orbit searches for P-BH binaries with spin-period $P_{\rm spin} \geq 13\,{\rm ms}$ covering the 3-10 $\mathrm{T_{obs}}$ regime are feasible for the High Time Resolution Universe South Lowlat latitude survey. Additionally, an elliptical orbit search for $P_{\rm spin} \geq 10\,{\rm ms}$ pulsars orbiting intermediate mass black holes in Globular clusters in the 5-10 $\mathrm{T_{obs}}$ regime is feasible for observations shorter than 2 hours with an eccentricity limit of 0.1.

5.1 Introduction

The discovery of the first binary pulsar PSR B1913+16 and the subsequent measurement of its orbital period decay provided a new and exciting tool for testing theories of gravity in the strong field limit (Hulse & Taylor, 1975b; Taylor & Weisberg, 1982; Weisberg & Taylor, 1984; Weisberg & Huang, 2016). Another major milestone was the discovery of the first, and to date only, known double pulsar system PSR J0737-3039A/B (Burgay et al., 2003; Lyne et al., 2004). Subsequent timing observations of this pulsar enabled measurement of seven Post-Keplerian parameters which provided multiple stringent tests of Einstein's theory of General Relativity (GR) (Kramer et al., 2006, 2021). Additionally, timing observations of the first pulsar triple system PSR J0337+1715 (Archibald et al., 2018; Voisin et al., 2020) set a new lower limit for the parameter Δ describing a possible violation of the Strong Equivalence Principle (SEP), consistent with GR and tightly constraining multiple scalar-tensor theories of gravity. These examples are only some of the highlights that demonstrate the value of discovering new relativistic binary pulsars. The anticipated discovery of a pulsar orbiting a black hole will provide a unique tool for probing strong-field gravity and black hole physics. Wex & Kopeikin (1999); Liu et al. (2014) gave a detailed account of how the properties of the black hole that can be studied by timing a pulsar orbiting the black hole. For example, the mass of a black hole can be uniquely determined by measuring two Post-Kepelerian parameters. Additionally, the spin of a black hole can be measured by studying the precession of the pulsar orbit with time caused by relativistic spin-orbit coupling. Similarly, the quadrupole moment of the black hole's external gravitational field can be detected with upcoming facilities like the Square Kilometre Array (SKA) (Kramer, 2004). Liu et al. (2014) demonstrated that by timing a pulsar orbiting a stellar-mass black hole (SBH) for few years would allow precise measurements of the black hole mass and spin. Such a discovery would help us test the Cosmic Censorship Conjecture and the “No-hair" theorem. Despite several observational efforts to detect such systems, none have been found so far. One explanation for this maybe the extreme Doppler smearing caused as the pulsar moves around its orbit during the observation time. The classical approach to account for this is do an acceleration search (Johnston & Kulkarni, 1991) i.e to first order assume that the acceleration of the binary pulsar is constant during the observation. This assumption holds weight only if the observation samples a small fraction of the orbit typically $P_{\rm orb} \leq 10\ T_{\rm obs}$ (Ransom et al., 2003; Ng et al., 2015). Therefore, in order to be sensitive to highly relativistic systems we would require very short observations. This however increases the minimum flux density of the pulsar that we can detect as sensitivity grows with ${T_{\rm obs}}^{1/2}$. Therefore, by using an acceleration search we have a trade-off between sensitivity and the ability to detect short-orbital period binaries. Another option is to do a higher-order polynomial search, for example a jerk search which assumes a constant jerk during the observation. Andersen & Ransom (2018) showed that using this technique leads to an increased sensitivity to short orbital period binaries $P_{\rm orb} \approx$ 5-15 $T_{\rm obs}$. Eatough et al. (2013) also demonstrated the improvements of jerk searches compared to acceleration searches for observations of PSR J0737-3039A. Smith (2016) proposed four new algorithms for radio pulsar searches including a semicoherent search which divides the data into small chunks and combines information from coherent subsearches while preserving as much phase information as possible. Yet another approach is do a fully coherent search on the full length observation by searching directly over the Keplerian parameters. Assuming a circular orbit binary, this would lead to a three-parameter search over orbital period, projected semi-major axis and initial orbital phase. Knispel (2011); Allen et al. (2013) and Knispel et al. (2013) did exactly this by using a matched filtering process, convolving the data with circular orbit templates in order to find pulsars in Pulsar Arecibo L-band Feed Array (PALFA; Cordes et al. 2006) and Parkes Multi-Beam Pulsar Survey (PMPS; Manchester et al. 2001) survey using the volunteer-distributed Einstein@Home project. Assuming standard binary stellar evolution model (Yungelson & Portegies Zwart, 1998; Voss & Tauris, 2003), PSR-BH binaries are expected to be in highly eccentric orbits (Shao & Li, 2018). While, we are already computationally limited with the acceleration and jerk search techniques, a yet unexplored question is how feasible would a fully coherent elliptical orbit search (five-parameter search) be and if there are any advantages in terms of sensitivity towards unexplored parameter spaces that a Keplerian search would provide. Recent advances in GPU technologies have considerably sped up the typically computationally expensive parts of pulsar search pipelines like dedisperison, time/frequency domain resampling, calculating a Fast-Fourier transform (FFT) and harmonic summing. Here, we present a GPU implementation of a five-parameter search pipeline using the template-bank algorithm that is more sensitive than acceleration and jerk search pipelines to PSR-SBH binaries in eccentric orbits. We do a detailed comparison of the detectability of mildly eccentric PSR-SBH binaries using our pipeline and compare to all of the currently used standard binary pulsar pipelines including those mentioned above. Finally, we quantify the computational feasibility of using this method for targeted and wide-area galactic plane observations in the near future.

5.2 Template-Bank Search

5.2.1 Signal Model and Detection Statistic

To discover new pulsars buried in our observations, we need to first define a signal model. We make the following assumptions in our signal model. The orbit of the pulsar is Keplerian.[1] We know the exact location of the source on the sky, therefore Doppler phase drifts caused due to the detector motion can be removed and the signal is monochromatic at the spin-frequency of the pulsar f_0 i.e we ignore spin-down effects of the pulsar. Therefore, the rotational phase Φ for the fundamental mode of the signal emitted by a uniformly rotating pulsar in an eccentric orbit observed from solar-system barycenter is

$$\Phi(t;\Lambda) = 2\pi f_0 t + \phi_D(t), \tag{5.1}$$

where t is the barycentric time, f_0 is the spin-frequency of the pulsar, Λ is a tuple of the unknown binary parameters of the pulsar signal and $\phi_D(t)$ is the Doppler phase correction factor due to the pulsar being in an eccentric orbit, given by (see section 2B of Dhurandhar & Vecchio (2001) for the full derivation),

$$\phi_D(t;\Lambda) = \frac{2\pi f_0 a \sin\varepsilon}{c}\left[\sin\psi\cos E(t) + \cos\psi\sqrt{1-e^2}\sin E(t)\right]. \tag{5.2}$$

Here, $E(t)$ is the eccentric anomaly which is a function of time, a is the semi-major axis of the orbit, ε is the inclination angle and $\frac{asin\varepsilon}{c}$ is the projected semi-major axis of the orbit in light-seconds. ψ is the longitude of periastron and e is the eccentricity of the orbit. The eccentric anomaly $E(t)$ is related to the mean angular velocity Ω and the mean anomaly $M(t)$ by the Kepler equation,

$$E(t) - e\sin E(t) = \Omega t + \alpha \approx M(t), \tag{5.3}$$

where α is the initial orbital phase of the pulsar orbit. We have seven unknown parameters $\Lambda = \{f_0, a, \varepsilon, \Omega, \alpha, \psi, e\}$ in equation 5.1. The semi-major axis of the pulsar cannot be directly observed. Since we search over different waveforms in our signal model, what matters for our search is the combination of parameters that affect the rotational phase, in this case the projected semi-major axis $\tau = \frac{a\sin\varepsilon}{c}$. Therefore, in practice a blind coherent search for binary pulsars in a Keplerian orbit needs to search over the six-dimensional parameter space of $\Lambda = \{f_0, \tau, \Omega, \alpha, \psi, e\}$. It is unlikely that without any prior information, we would exactly sample the true parameters of the undiscovered pulsar, therefore we use a matched filtering process of convolving the data with multiple waveforms. Each sample trial consists of different six-tuple parameter combinations and is called a template of the signal. A combination of many such templates form a template bank. The density of templates in the template-bank can be characterised by its mismatch value (defined in equation 5.12) i.e the worst-case detec-

---

[1] This assumption is based on the typical timescales of search-mode observation of radio pulsars which is in the order of few hours. For long observations lasting months or years, additional parameters including Post-Keplerian and spin-down parameters need to be added to the phase model.

tion statistic loss. Our general goal is to cover large parts of the parameter space with the minimum number of templates possible in order to minimise computational costs. This is a widely used technique especially in gravitational-wave analysis, see Owen & Sathyaprakash (1999); Abbot et al. (2007, 2009). In order to sample in this parameter space efficiently, we need to define distances between two points. This is given by the parameter space metric (Owen, 1996) defined in the next section. For completeness we also describe the signal phase assuming a circular orbit. In this case eccentricity is zero and the longitude of periastron and the eccentric anomaly are undefined, therefore the phase model becomes

$$\phi(t;\Lambda) = 2\pi f_0\big[t + \tau\sin(\Omega t + \alpha)\big]. \tag{5.4}$$

In this case we search over the four-dimensional parameter space of $\Lambda = \{f_0, \tau, \Omega, \alpha\}$.

The total time-domain radio intensity signal is a sum of instrumental and environmental noise $N(t)$ and harmonic sums of a pulsar signal.

$$s(t;\Lambda) \equiv N(t) + \sum_{n=1}^{\infty} s_n(t;\Lambda), \tag{5.5}$$

where the intensity of each harmonic is given by

$$s_n(t;\Lambda) \equiv \Re\big[\mathcal{A}_n \exp[-in\Phi(t;\Lambda)]\big], \tag{5.6}$$

where $\mathcal{A}_n$ are the complex amplitudes of the harmonics of the signal. We then define a coherent detection statistic $P_n(\Lambda, \Lambda')$ for the nth harmonic computed using the radio intensity correlated with the nth normalised signal template currently being searched $\exp[-in\Phi(t;\Lambda)]$. This is the detection statistic recovered from a pulsar signal with true parameters at Λ with a template at Λ'.

$$P_n(\Lambda, \Lambda') = \left|\frac{1}{T}\int_0^T dt\, s(t;\Lambda)\exp\big[-in\Phi(t;\Lambda')\big]\right|^2. \tag{5.7}$$

Thresholding the detection statistic helps us minimize the false-alarm probability at a fixed rate of false-alarm probability (Allen et al., 2002). In the absence of a pulsar signal, the $N(t)$ term dominates and the detection statistic term is proportional to the instrumental noise whereas in the presence of a strong signal the $N(t)$ term can be neglected and the expectation value of the detection statistic then becomes,

$$\langle P_n(\Lambda, \Lambda')\rangle \approx \left|\frac{A_n}{2}\right|^2 \left|\frac{1}{T}\int_0^T dt\, \exp[in(\Phi(t;\Lambda) - \Phi(t;\Lambda'))]\right|^2. \tag{5.8}$$

While searching for new pulsars, since we do not known a *priori* the pulse-profile, we assume the Fourier transform of the pulse profile resembles a Dirac delta function truncated to a finite number of harmonics. Therefore, we equally weight P_n for different harmonics (P_0,..P_4) and sum them together to form our combined detection statistic.

$$S_L = \sum_{n=1}^{2^L} P_n. \tag{5.9}$$

where 2^L is the number of harmonic sums performed. In the noise-only case, assuming Gaussian statistics, the probability distribution function (PDF) of S_L is a χ^2 distribution with 2N = 2^{L+1} degrees of freedom. Integrating this PDF shows that the probability for the power in any spectral bin to exceed a threshold $P_{\rm min}$ is proportional to exp(-$P_{\rm min}$). This is called the false-alarm probability i.e the chance that a candidate arises from random noise fluctuations rather than a true pulsar signal. $P_{\rm min}$ can be estimated by setting the number of false postive to one.

$$P_{\rm min} = -\ln\left(1/(2n_{\rm samp})\right), \tag{5.10}$$

where $n_{\rm samp}$ is the number of samples in the timeseries. Lorimer & Kramer (2012) showed that we can convert the power threshold $P_{\rm min}$ into a Signal to Noise ratio threshold S/N$_{\rm min}$ given by

$$\mathrm{S/N}_{\rm min} = \frac{\sqrt{\ln\left[\mathrm{n_{trials}}\right]} - \sqrt{\pi/4}}{\sqrt{1-\pi/4}}, \tag{5.11}$$

where $n_{\rm trials} = n_{\rm samp} \times n_{\rm DM\,trials} \times n_{\rm harmonic\,sums} \times n_{\rm orbital\,trials}$. In practise, due to the presence of Radio frequency interference (RFI) a higher threshold maybe required. We consider all candidates with Signal to Noise ratio greater than S/N$_{\rm min}$ to be statistically significant and visually inspect them.

5.2.2 Definition of Metric and Mismatch

Using equation 5.8, we can define the mismatch (fractional loss of detection statistic, in our case signal to noise ratio) between two points as

$$m(\Lambda, \Lambda') = 1 - \frac{P_n(\Lambda, \Lambda')}{P_n(\Lambda, \Lambda)} \approx g_{\alpha\beta}\, \Delta\Lambda^{\alpha} \Delta\Lambda^{\beta} + \mathcal{O}(\Delta\Lambda^3), \tag{5.12}$$

where a mismatch of one implies complete loss of signal and a mismatch of zero implies a perfect recovery of the signal, $g_{\alpha\beta}$ is the metric tensor, α and β correspond to points in the six-parameter space described earlier and we adopt the Einstein summation convention where repeated indices are summed over. For small deviations of the parameter space coordinates, $g_{\alpha\beta}$ can be calculated by Taylor expansion as (section 8.5 of Knispel (2011) for the full derivation)

$$g_{\alpha\beta} = \langle \partial_\alpha \Phi \partial_\beta \Phi \rangle_T - \langle \partial_\alpha \Phi \rangle_T \langle \partial_\beta \Phi \rangle_T, \tag{5.13}$$

where the angle brackets represent time average of a function $G(t)$

$$\langle G(t)\rangle_T \equiv \frac{1}{T}\int_0^T dt G(t). \tag{5.14}$$

For computing the time-averages of the derivatives of the phase with respect to the search parameters, it is convenient to express $\cos E$ and $\sin E$ as a power series in the eccentricity parameter e with harmonics in the mean anomaly M

$$\cos E = \sum_{k=0}^{\infty} C_k(e)\cos(kM), \tag{5.15}$$

$$\sqrt{1-e^2}\sin E = \sum_{k=1}^{\infty} S_k(e)\sin(kM), \tag{5.16}$$

where C_k and S_k are power series in e. The order to which one should consider expanding the power series depends on the spin-frequency and eccentricity we are interested in searching for. See Appendix C of Nieder et al. (2020) for a derivation of this power series. In this work we consider up-to the 7th power in e, and these values can be found in Appendix E.

5.2.3 Random Template Banks

The central idea behind this algorithm is to give up the requirement of complete coverage of the parameter space ($\eta = 1$). Instead, we aim for a user-defined probabilistic coverage of the parameter space ($\eta < 1$). The templates are distributed randomly with uniform probability based on values of the square root of the determinant of the parameter space metric. We refer the reader to Messenger et al. (2009) for an in-depth review of this algorithm including derivations for the expressions mentioned here. Here we summarise the relevant details that are applicable to our work. Using the definition of mismatch (described in equation 5.12), we can define a metric for our parameter space $g_{\alpha\beta}$ for different signal parameter values $(\lambda_1, \lambda_2, ..., \lambda_n)$ and calculate the proper volume V_{S_n} of our parameter space S_n by:

$$V_{S_n} = \int_{S_n} dV = \int_{S_n} d^n\lambda\sqrt{g}, \tag{5.17}$$

where $g = \det(g_{\alpha\beta})$ is the determinant of the parameter space metric. We then define the volume of an n-dimensional ball with unit radius as:

$$C_n = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)}, \tag{5.18}$$

This is the n-dimensional volume enclosed by a (n-1) dimensional sphere. The volume covered by a single template with mismatch m_* is:

$$V_T = C_n m_*^{\frac{n}{2}}. \tag{5.19}$$

The total number of random templates required to achieve a coverage η with nominal mismatch m_* of a parameter space S_n is thus:

$$N_R(\eta, m_*, S_n) = \frac{\ln(1-\eta)}{\ln\left(1 - m_*^{\frac{n}{2}} \frac{C_n}{V_{S_n}}\right)}. \tag{5.20}$$

In most cases, we are interested in large parameter spaces where the volume of the parameter space V_{S_n} is much greater than the volume of a single template V_T, therefore, $m_*^{\frac{n}{2}} \frac{C_n}{V_{S_n}} \ll 1$, and we can Taylor-expand equation 5.20 to write,

$$N_R(\eta, m_*, S_n) \approx \frac{m_*^{\frac{-n}{2}} V_{S_n}}{C_n} \ln\left(\frac{1}{1-\eta}\right). \tag{5.21}$$

5.2.4 Frequency-Projected Metric

Our template bank construction method is similar to previous works by Knispel (2011); Allen et al. (2013) and Knispel et al. (2013). We do this by computing a Cartesian product of a five-dimensional orbital template bank $\Lambda_{\text{orb}} = \{\tau, \Omega, \alpha, \psi, e\}$ along with a uniformly spaced grid in the spin-frequency axis. This is mainly done for computational reasons as the latter can be efficiently calculated using the FFT algorithm (Cooley & Tukey, 1965). Therefore, we take the metric tensor (defined in equation 5.13) and project it to a subspace of constant spin frequency f_0. This can be thought of as a slice across the parameter-space for a given f_0. This is usually chosen to be the highest spin-frequency of the harmonic of the signal we would like to detect in our search $f_{\max}$. This number should be chosen carefully as the number of templates for an five-dimensional coherent elliptical orbit search grows to the fifth power of spin frequency $\left(\mathrm{N}_{\text{templates}} \propto \mathrm{f}_0^5\right)$. Choosing a particular $f_{\max}$ does not automatically imply that we cannot detect signals with $f_0 > f_{\max}$. This usually means that our mismatch values would be higher for higher spin frequencies and conversely our mismatch values will be lower for lower spin frequencies. In other words, this means that we over-sample our parameter space for lower spin frequency values, under-sample for higher spin-frequency values and we are optimally sampled for the $f_{\max}$ we choose to create our template-bank. We calculate the frequency projected metric as

$$\gamma_{\alpha\beta} = g_{\alpha\beta} - \frac{g_{f\alpha} g_{f\beta}}{g_{ff}}, \tag{5.22}$$

where we the repeated indices are summed over. Therefore, for each orbital template, we apply a time-domain resampling algorithm (described in Section 5.5) and then calculate an FFT.

5.2.5 Signal Phase and Dimensionless Parameters

Following the formalism described in Dhurandhar & Vecchio (2001), we re-write our signal phase in dimensionless parameters in order to simplify the calculation for the

Table 5.1: Prior ranges used for generating our orbital template bank. U here refers to a uniform probability distribution.

Parameter	Value/Range	Unit
t_{obs}	1.2	h
$\text{m}_{\text{pulsar,min}}$	1.4	$M_\odot$
$\text{m}_{\text{companion,max}}$	8.0	$M_\odot$
Coverage (η)	0.9	
Mismatch (m)	0.2	
Max.Spin Frequency (f_{max})	66.67	Hz
Orbital Period (P_{orb})	U(6, 12)	h
Projected Radius (τ)	$\text{U}\left(0, \frac{G^{\frac{1}{3}}\Omega_{\text{orb}}^{-\frac{2}{3}} m_{\text{comp,max}}}{c\left(m_{\text{pulsar,min}} + m_{\text{comp,max}}\right)^{\frac{2}{3}}}\right)$	lt-s
Initial Orbital Phase (α)	U(0, 2 π)	rad
Longtitude of Periastron (ψ)	U(0, 2 π)	rad
Eccentricity (e)	U(0, 0.1)	

determinant of the parameter space metric. If we multiply our signal phase model Φ by a constant factor χ, such that

$$\Phi = \chi\tilde{\Phi}, \tag{5.23}$$

then each component of the metric $g_{\alpha\beta}$ and the frequency-projected metric $\gamma_{\alpha\beta}$ become

$$g_{\alpha\beta} = \chi^2 \tilde{g}_{\alpha\beta}, \quad \gamma_{\alpha\beta} = \chi^2 \tilde{\gamma}_{\alpha\beta}, \tag{5.24}$$

and the determinants are scaled by

$$det(g_{\alpha\beta}) = \chi^{2N+2} \det(\tilde{g}_{\alpha\beta}), \quad det(\gamma_{\alpha\beta}) = \chi^{2N} \det(\tilde{\gamma}_{\alpha\beta}) \tag{5.25}$$

As shown in equation 5.17 and 5.21, the proper volume and the total number of templates is proportional to the square root of the determinant of the metric tensor, which in these coordinates becomes,

$$V_{S_n} = \chi^N \tilde{V}_{S_n}, \quad N_R = \chi^N \tilde{N}_R. \tag{5.26}$$

where the 'N' in the exponent refers to the number of free parameters in our signal model. In the circular case N=3 and in the elliptical case N=5. Using these properties, we re-write our signal phase model as

$$\Phi = (2\pi f_0 T)\tilde{\Phi}, \tag{5.27}$$

where

$$\tilde{\Phi} = \kappa u + X\cos E + Y\sqrt{1-e^2}\sin E. \tag{5.28}$$

The dimensionless parameters are defined as,

$$\begin{aligned} u &= \frac{t}{T}, \\ \kappa &= \frac{f - f_0}{f_0}, \\ X &= \frac{\tau \sin\psi}{T}, \\ Y &= \frac{\tau \cos\psi}{T}, \\ \Omega' &= \Omega T. \end{aligned} \tag{5.29}$$

Here u is dimensionless time, satisfying $0 \leq u \leq 1$, κ is dimensionless frequency satisfying $-1 \leq \kappa \leq 0$. Ω' is the number of orbits in radians and $\frac{\Omega'}{2\pi}$ is the number of orbits covered by the binary during the observation time T. From equation 5.26 and 5.27, it is clear that

$$\chi = 2\pi f_{\max} T, \tag{5.30}$$

where we substitute $f_0 = f_{\max}$ which is the maximum spin-frequency of the pulsar we want our search to be sensitive to with the chosen mismatch. Using the dimensionless rotational phase (defined in equation 5.28) we then proceed to calculate the metric tensor $g_{\alpha\beta}$ (using equation 5.13) with parameters $\Lambda = \{\kappa, X, Y, e, \Omega, \alpha\}$ and frequency-projected metric $\gamma_{\alpha\beta}$ with parameters $\Lambda = \{X, Y, e, \Omega, \alpha\}$. We then calculate the square root of the determinant of this frequency projected metric tensor. Finally, we multiply this value by our scaling factor $(2\pi f_{\max} T)^5$.

5.2.6 Chosen Parameter-Space

Before we start constructing a template bank to search for binary pulsars, we first need to decide what region of the parameter space will we focus our search on. The two main factors that are important for this are our astrophysical motivation and computational feasibility. The former is based on our prior information about binary stellar evolution and past searches on our data which can help us exclude regions that are unlikely to yield promising results or already been investigated. The most important criteria is how much computational resource is available and how much time can be invested in doing our search analysis. With unlimited computing power we could explore all regions of our parameter space, however in practice we try to balance both these factors by maximising our chances of detecting an exciting binary pulsar at a fixed computational cost.

The most commonly used search techniques to find binary pulsars in radio observations are acceleration searches which can done in both the time domain (Johnston & Kulkarni, 1991) and frequency domain (Ransom et al., 2002). Recently, PRESTO[2] has added support for frequency-domain jerk searches (Andersen & Ransom, 2018). Allen et al. (2013); Knispel et al. (2013) did a coherent search using the template-bank

---

[2] `https://github.com/scottransom/presto`

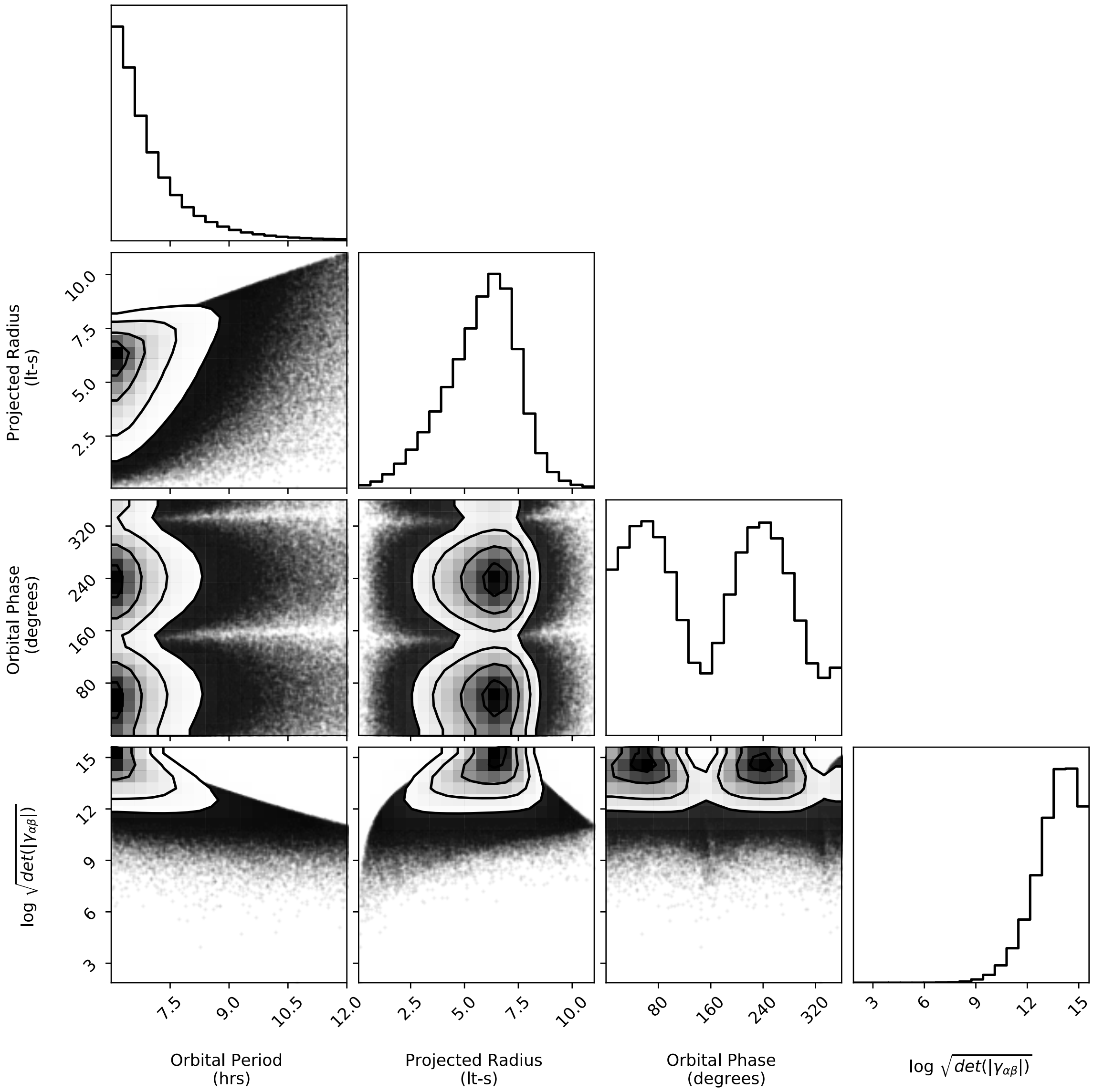


Figure 5.1: Corner plot showing the distribution of the value of $\sqrt{|\gamma_{\alpha\beta}|}$ assuming a circular orbit which is a measure of our template pseudo-density for different orbital parameter combinations. Black points show the relative sampling density in the template bank. Histograms in the main diagonal show the one-dimensional marginalized distribution of the template density across each of our orbital parameters. We see that there is exponential growth of orbital templates for short orbital periods and high $a \sin \varepsilon$ values. Contours in the 2-D plots highlight values lying with 0.5, 1, 1.5 and 2 sigma respectively. $a \sin \varepsilon = 0$ is the special case of face-on orbit when there is no Doppler modulation of the signal. In the orbital phase versus orbital period plot, we see two streaks of low values of $\sqrt{|\gamma_{\alpha\beta}|}$. This corresponds to $\left(1 - \frac{T_{\rm obs}}{P_{\rm orb}}\right) \times 180$ and $\left(2 - \frac{T_{\rm obs}}{P_{\rm orb}}\right) \times 180$ degrees. At these orbital phases, the pulsar is either approaching towards or receding from the observer essentially along the line of sight, and the determinant tends to zero. This makes the signal appear Doppler shifted by a constant, depending on the velocity of the source that does not need to be corrected for.

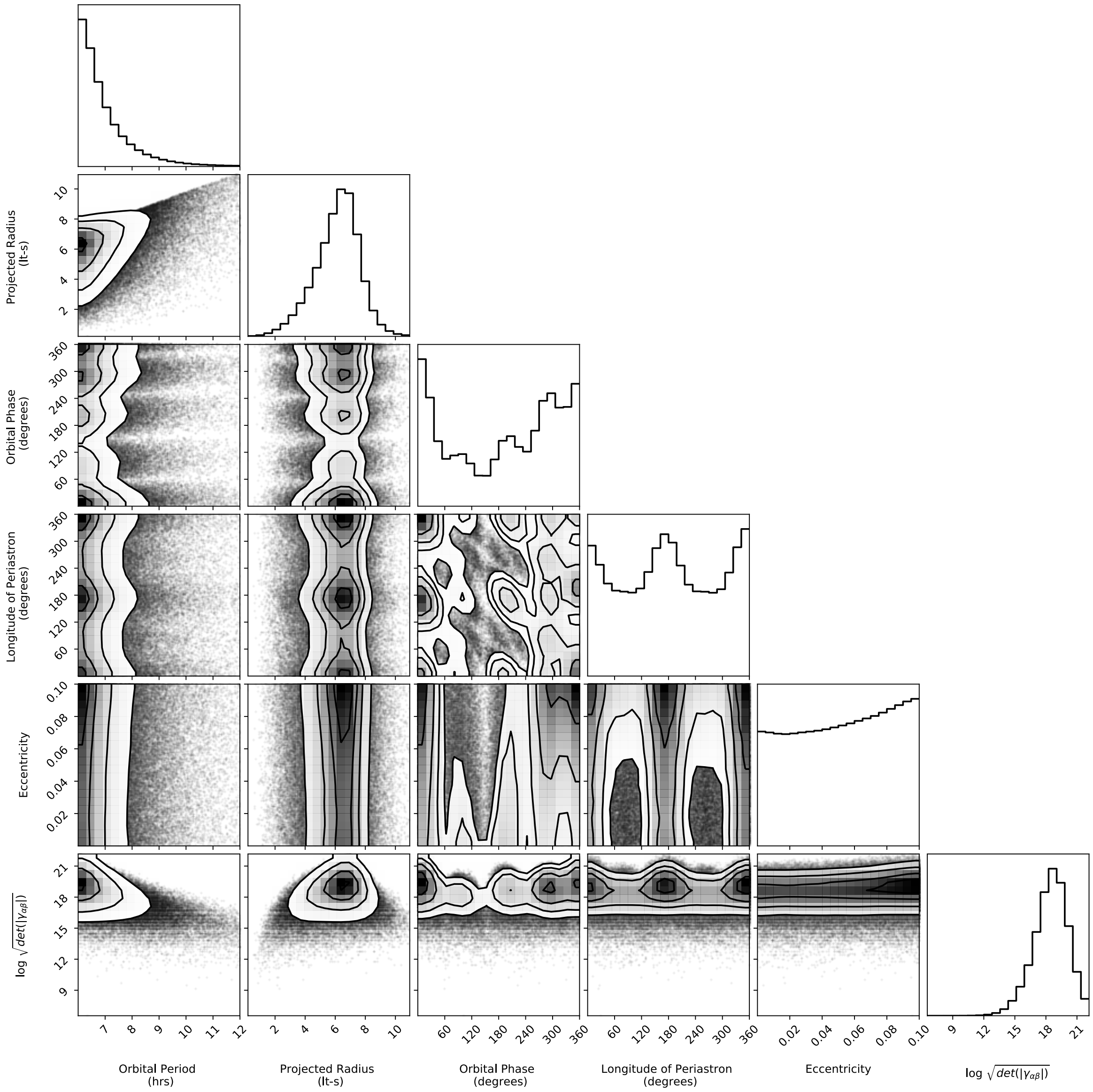


Figure 5.2: Corner plot showing the distribution of the value of $\sqrt{|\gamma_{\alpha\beta}|}$ which is a measure of our template pseudo-density assuming an elliptical orbit for different orbital parameter combinations. Black points show the relative sampling density in the template bank. As in the circular case, there is exponential increase in the amount of required orbital templates for a combination of short orbital periods and high $a \sin \varepsilon$ values. Additionally, we notice a quadratic dependence of templates across the eccentricity parameter and a cosine dependence across the longitude of periastron parameter. Contours in the 2-D plots highlight values lying with 0.5, 1, 1.5 and 2 sigma respectively.

Table 5.2: Input, Spin and Orbital Parameters for the Simulated Pulsar-Black hole binaries used in our analysis. We simulated all our orbits with a Dispersion Measure DM = 0. Here, 'U' denotes a uniform distribution.

Parameter	Value/Range	Unit
t_{obs}	1.2	h
$\mathrm{m}_{\mathrm{pulsar,min}}$	1.4	$M_{\odot}$
$\mathrm{m}_{\mathrm{companion,max}}$	8.0	$M_{\odot}$
Spin-Period	15.0	ms
Duty Cycle	10.0	%
Dispersion Measure	0.0	$\mathrm{pc\,cm^{-3}}$
Frequency Channels	1024	
Bandwidth	400	MHz
Frequency of Channel1	1181.804688	MHz
Channel bandwidth	0.390625	MHz
Signal-to-noise ratio of single pulse	0.02	
Data Bit-Rate	4	
Number of samples	2^{24}	samples
Sampling time	256e-06	s
Orbital Period ($\mathrm{P}_{\mathrm{orb}}$)	U(6, 12)	h
Eccentricity (e)	0.1	
Initial Orbital Phase (α)	U(0, 2 π)	rad
Longtitude of Periastron (ψ)	U(0, 2 π)	rad
Projected Radius (τ)	$\mathrm{U}\left(0, \dfrac{G^{\frac{1}{3}}\Omega_{\mathrm{orb}}^{-\frac{2}{3}} m_{\mathrm{comp,max}}}{c\,(m_{\mathrm{pulsar,min}} + m_{\mathrm{comp,max}})^{\frac{2}{3}}}\right)$	lt-s

algorithm described earlier (assuming circular orbits - three orbital search parameters) to search for compact binaries within PALFA (Cordes et al., 2006) and PMPS (Manchester et al., 2001) observations respectively. Given the extensive amount of work that have already been done building and testing these software suites and their remarkable success in finding several relativistic binaries like PSR J0737-3039A (Burgay et al., 2003) which was discovered in a FFT search, PSR J1757-1854 (Cameron et al., 2018) which was discovered in a segmented acceleration search among others, our goal here is to complement their work by focusing our search on regions of parameter space that were not searchable by these algorithms. In Section 5.3.1 we do a detailed simulation comparing these search algorithms. We fix our upper limit for orbital period as $P_{\text{orb}} = 10T$. This is because the assumption of "constant" acceleration works well for long orbital period binaries i.e when $P_{\text{orb}} > 10T$. The lower limit of the orbital period searched is determined solely by the amount of computational power available. The constraint on projected semi-major axis, $a \sin \varepsilon$, is calculated based on the orbital period, masses of the pulsar and companion, and the different inclination angles we want to be sensitive to. Using Kepler's third law, we define

$$0 \leq a \sin \varepsilon \leq \theta \frac{G^{\frac{1}{3}} \Omega_{\text{orb}}^{-\frac{2}{3}} m_{\text{comp,max}}}{c \left(m_{\text{pulsar,min}} + m_{\text{comp,max}}\right)^{\frac{2}{3}}} \tag{5.31}$$

where $m_{\text{comp,max}}$ is the maximum companion mass, $m_{\text{pulsar,min}}$ is the minimum pulsar mass, G is the gravitational constant, c is the speed of light in vacuum, the parameter θ $(0 \leq \theta \leq 1)$ constrains the inclination angle that we want to be sensitive towards for a given binary system. In our template bank, we selected the following binary parameter ranges: $m_{\text{pulsar,min}} = 1.4\, M_{\odot}, m_{\text{comp,max}} = 8.0\, M_{\odot}$, $\theta = 1, P_{\text{orb}} =$ 6-12 hrs. We had no constrains on the initial orbital phase α ($0 \leq \alpha \leq 2\pi$) and longitude of periastron ψ $(0 \leq \psi \leq 2\pi)$. The upper-limit on the eccentricity range is also decided by computational factors. We limit our searches to a range between $0 \leq e \leq 0.1$. The prior ranges used in our tests are summarised in table 5.1.

5.2.7 Distributing templates in the parameter space

Once we have calculated the number of required orbital templates using equation 5.21, the next step is to generate the template-bank, i.e. distribute the orbital templates in the parameter space. Our goal is to sample from a probability density function that is proportional to the square root of the determinant of the metric tensor $\sqrt{|\gamma_{\alpha\beta}|}$ multiplied by the prior probabilities in our orbital parameter range. Markov Chain Monte Carlo (MCMC) methods are well suited for such problems. We use the PYTHON package EMCEE (Foreman-Mackey et al., 2013) which is an implementation of the affine-invariant MCMC ensemble sampler proposed by (Goodman & Weare, 2010). Briefly, this method works as follows. We explore the parameter space through a set of "walkers" which run in parallel. Each walker represents a point in the parameter space. After each iteration, a walker selects another walker and takes a step along the line in parameter space connecting between them. These step sizes are chosen stochastically. At this new step, we then calculate our posterior distribution function (PDF). We always accept steps where the PDF value increases and sometimes accept steps if the PDF value decreases. These walkers can be thought of as separate Metropolis-Hastings chains running in parallel with the caveat that the steps taken by one walker are dependent on the position of the other walkers. Based on the recommendation in Goodman & Weare (2010), we use the integrated autocorrleation time to quantify the effects of sampling error in our chain. We assume our chains have converged when the length of the chain is a hundred times our integrated autocorrleation time averaged across all dimensions. We chose 800 walkers for our chain, and once the chains have sufficiently converged, we discard the burn-in samples and randomly select the number of required templates from our chain. Using this method, we generated two orbital template banks, one assuming a circular orbit and another assuming an elliptical orbit. Figure 5.1 shows a corner plot (Foreman-Mackey, 2016) for the circular orbit case. Plots along the longest diagonal show the one dimensional marginalized distribution of the template pseudo density for each of the orbital parameters. The off-diagonal plots show the marginalized two dimensional distributions for different combinations of the orbital parameters. We see that the number of templates grow exponentially for shorter orbital periods and a combination of short orbital periods and high-$a\sin\varepsilon$ values. We see a dip in the template density for high values of $a\sin\varepsilon$ because longer orbital period binaries tend to higher $a\sin\varepsilon$ value ranges due to kepler's third law. In the orbital phase versus orbital period plot, we see two streaks of very low values of $\sqrt{|\gamma_{\alpha\beta}|}$. This corresponds to $\left(1 - \frac{T_{\rm obs}}{P_{\rm orb}}\right) \times 180$ and $\left(2 - \frac{T_{\rm obs}}{P_{\rm orb}}\right) \times 180$ degrees. At these orbital phases, the pulsar is either approaching or receding from the observer essentially along the line of sight, and the determinant tends to zero. This makes the signal appear Doppler shifted by a constant, depending on the velocity of the source that does not need to be corrected for.

Figure 5.2 shows the corner plot for the elliptical orbit case. As in the circular case, the plots along the diagonal show the one-dimensional marginalized distribution of $\sqrt{|\gamma_{\alpha\beta}|}$ for each of the five orbital parameters. We see that the overall value of $\sqrt{|\gamma_{\alpha\beta}|}$ has increased by three orders of magnitude. We see a similar increase in the number of trials for short orbital periods and a combination of short orbital periods and high $a\sin\varepsilon$ values. There is a quadratic dependence of the template density across the eccentricity parameter. The sinusoidal dependence of the density of templates across orbital phase in the circular orbit case has been washed away in this plot, due to the different eccentricity values in our simulations. Additionally, we also notice a near cosine dependence of template density across the longitude of periastron parameter.

Using the metric approximation defined in section 5.2.2 leads to an over-estimation of the required orbital templates in the template bank. One method to reduce the total number of orbital templates is to apply the Stochastic template bank algorithm (Harry et al., 2009). Stochastic template banks are formed in a similar manner with an additional second step that prunes orbital templates that are closer than the nominal mismatch value. We provide an implementation of this algorithm in our software suite. Creating a stochastic template bank is computationally expensive. Therefore, the trade-off between calculating a stochastic template bank and time-saved from search analysis should be evaluated. For all our tests in the subsequent sections we use a random template bank to save compute time.

5.2.8 Template Bank Verification

After constructing the template bank, the next step is to verify that the template bank satisfies the chosen coverage and mismatch criteria. We did this by simulating PSR-SBH binaries in a 72-minute observation with the pulsar rotating at the maximum spin-period chosen for our template bank $P_{\mathrm{spin}} = 15$ ms. We simulated 10,000 binaries with a uniform distribution of orbital period between 6-12 hours, a fixed eccentricity of 0.1 and inclination angle varying from 0 to 90 degrees. We let the orbital phase and longitude of periastron vary between a value from 0 to 2π radians. All the parameters used in the simulation can be found in table 5.2. We also simulated an isolated pulsar in order to calculate our best case signal to noise ratio. The mismatch distribution from running our elliptical binary search pipeline is shown in figure 5.3. We created the template bank to have a coverage of $\eta = 90\%$ and a mismatch of 0.2. Since we also search over frequency using an FFT, we need to account for the additional mismatch from this parameter. The response of the FFT in frequency is not uniform. It's maximum for signals falling in the centre of the fourier bin and loses sensitivity as you go towards the edge of a bin. This is called *scalloping* (see for e.g. Middleditch et al. (1993)). One option to mitigate this is by padding the timeseries with its mean value to minimise the effect of FFT scalloping (See section 3.8.4 of Knispel (2011) or Ransom et al. (2002) for a detailed discussion on this). We avoid this in our analysis in order to minimise computational time. The 90th percentile of our mismatch distribution is 0.33, i.e. we cover 90% of our parameter space with a mismatch value $m_{0.9} \leq 0.33$. Therefore, the requirements of the template-bank have been fulfilled for our chosen search set-up.

5.3 Comparing to other pulsar-search pipelines

5.3.1 Tests on Simulated Observations of Mildly Eccentric P-SBH Binaries

In this section, we compare the performance of a PRESTO -based pipeline doing an acceleration and jerk search to a template-bank search over three Keplerian parameters assuming a circular orbit and over five Keplerian parameters assuming an elliptical orbit. We restrict our analysis to binaries in the 5-10 T_{obs} regime. We start by simulating 10,000 PSR-SBH Binaries with an eccentricity of 0.1, companion mass of 8 $M_{\odot}$, pulsar mass of 1.4 $M_{\odot}$ and a spin-period of 15 ms. These values are kept fixed. We simulate orbits with a uniform distribution of orbital period between 6 and 12 hours along with a uniform distribution of inclination angles, orbital phase and longtitude of periastron values. We used a modified version of the code FAKE from SIGPROC[3] for our work. A

---

[3] `https://github.com/SixByNine/sigproc`

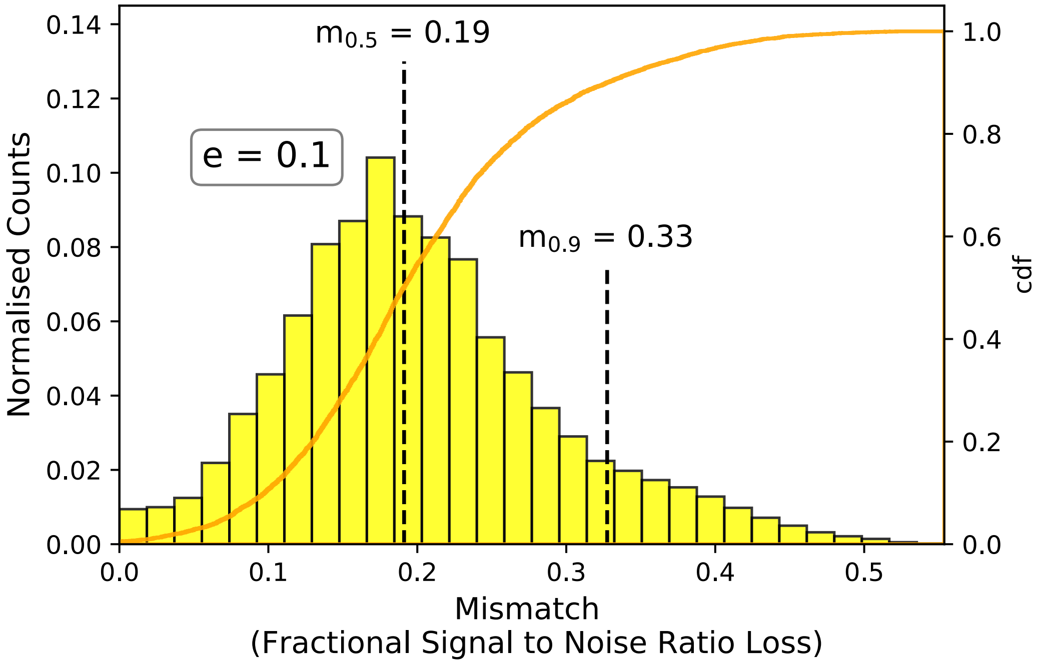


Figure 5.3: Mismatch distribution of our template-bank for simulated pulsar-black hole binaries. The dotted line indicates the median and 90th percentile mismatch values.

full list of our input spin and orbital parameters for the simulation can be found in table 5.2. Additionally, we simulated an isolated pulsar with identical input parameters in order to calculate the best-case detection significance that can be recovered. For each simulated binary pulsar, we run each of our pipelines and calculate the mismatch value using the following expression.

$$m = 1 - \frac{(\mathrm{SNR})_{\mathrm{Recovered,SearchPipeline}}}{(\mathrm{SNR})_{\mathrm{Isolated}}}. \tag{5.32}$$

PRESTO implements a frequency-domain acceleration/jerk search (FDAS/FDJS). FDAS assumes that the pulsar's acceleration A is constant throughout the observation. Under this assumption, each harmonic of the pulsar signal will experience an acceleration A given as:

$$A = \frac{\dot{f}c}{hf} = \frac{zc}{hf{t_{\mathrm{obs}}}^2}, \tag{5.33}$$

where z is the Fourier frequency bin number, h is the harmonic number of the signal, f is the spin-frequency and $\dot{f}$ is the spin-frequency derivative. Similarly, FDJS assumes a constant jerk, J, throughout the observation and the previous equation can be expanded to

$$\dot{A} = J = \frac{\ddot{f}c}{hf} = \frac{wc}{hf{t_{\mathrm{obs}}}^3}, \tag{5.34}$$

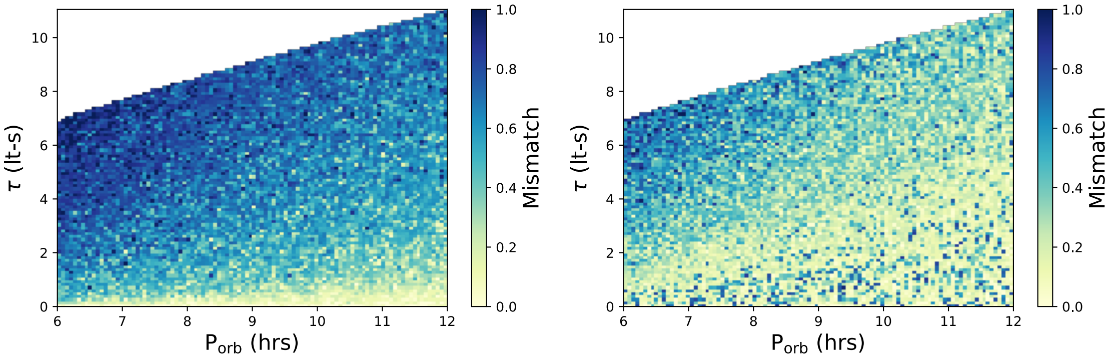
where w corresponds to the number of Fourier frequency derivative bins that the signal is smeared onto during the course of an observation. The user can input the maximum number of Fourier frequency bins ($z_{\max}$) and/or the maximum number of Fourier frequency derivative bins ($w_{\max}$) to search over in an observation.

For FDAS, we used the maximum value allowed by the software $z_{\max} = 1200$ which for a 15-ms pulsar in a 72-minute observation corresponds to an acceleration of 289.35 ms^{-2} for the fundamental ($h = 1$). For FDJS, we used a $z_{\max} = 600$ and a $w_{\max} = 1800$ which corresponds to an acceleration of 144.68 ms^{-2} and a jerk of 0.1 ms^{-3}. The maximum values used for acceleration and jerk in our simulations were limited by software and available RAM. In principle, these simulations could be repeated with a broader range of acceleration and jerk, however we do not expect a significant improvement in performance as the assumption of constant acceleration or jerk throughout the observation breaks down for extreme values.

For the template-bank search, we implemented two different searches. One assumes a circular orbit binary and searches over the three Keplerian parameters (P_{orb}, τ, α) and our second search expands the parameter space to include eccentricity e and longitude of periastron (ψ). For the circular search, we generate 10,128 orbital templates using the random template-bank algorithm with the same binary parameter range as our simulations but with $e = 0$. For the elliptical search, we generated 2,04,7716 orbital templates. Our results are shown in figure 5.4. These plots were made on a 100×100 grid. At each point, we simulated a binary pulsar and report the recovered mismatch value from each pipeline. The plots shows the mismatch value averaged across initial orbital phase and longitude of periastron. We notice a general trend of loss of signal-to-noise ratio (high mismatch values) using polynomial based searches for binaries with short orbital periods. Additionally, a high $a \sin \varepsilon$ value resulting from a steeper inclination angle also makes them harder to detect. For the special case of $a \sin \varepsilon = 0$ (face-on) orbit, the Doppler modulation is negligible, hence we notice a low mismatch value from all pipelines. Jerk search provides a significant improvement over acceleration search. However, this approach also loses sensitivity for binary pulsars with short orbital period binaries and high $a \sin \varepsilon$ values. The coherent search over three Keplerian parameters provides a uniform sensitivity throughout our parameter space with an average mismatch value of 0.35. Depending on the flux density of the pulsar, this value may or may not be acceptable. As expected, we notice the most improvement by searching over all five Keplerian parameters. We notice an average mismatch value of 0.19 for pulsars in our parameter space. The 1-D histogram for the mismatch values for all the pipelines can be seen in figure 5.5. Additionally, we also report the median and 90th percentile of our mismatch distribution in table 5.3.

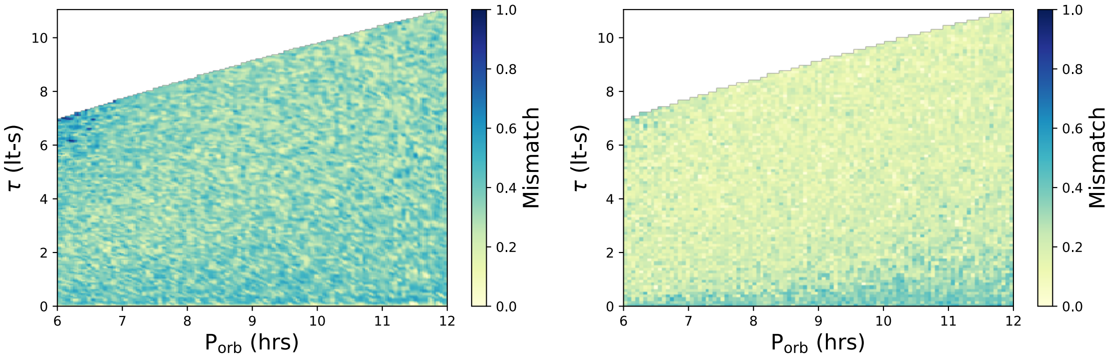
5.3.2 Tests on the Double PSR J0737-3039

In this section, we do a similar comparison of the performance of different pulsar search pipelines on an archival Parkes observation of the double pulsar PSR J0737-3039 that covers one full orbit of the system. The original observation had an integration time of 2.6 hours with a time sampling interval of 80 µs. The data was dedispersed at



(a) Acceleration Search over f and a ($z_{max} = 1200$)

(b) Jerk Search over f, a and j ($z_{max} = 600$, $w_{max} = 1800$)



(c) Template-Bank Search over f and Three Circular Orbit Keplerian Parameters.

(d) Template-Bank Search over f and Five Elliptical Orbit Keplerian Parameters.

Figure 5.4: Recovered signal-to-noise ratio expressed in terms of mismatch using different pulsar-search pipelines for simulated PSR-SBH binaries with eccentricity 0.1. Mismatch values shown here are averaged across different values of initial orbital phase and longitude of periastron. These plots contain a grid with 10,000 points. In each point, we simulate a binary, and report the recovered mismatch. As expected, we notice a high loss in signal to noise ratio for short orbital periods for an acceleration search. We notice a significant improvement with a jerk search, however it still fails to recover signals for short orbits and high asin ε values. We notice a further improvement by searching over three Keplerian parameters with an average mismatch value of 0.35 throughout the entire parameter space. The average mismatch value by searching over all five Keplerian parameters is 0.18. The cloudy structure in the template-bank search plots come from the fact that we have a probabilistic coverage of the parameter space $\eta = 0.9$, therefore for some regions we have a higher mismatch than the nominal value.

Table 5.3: Median and 90th percentile of mismatch values from our simulations for different pulsar search pipelines

Pipeline	m_{50}	m_{90}
Acceleration Search ($z_{max} = 1200$)	0.66	0.86
Jerk Search ($z_{max} = 600$, $w_{max} = 1800$)	0.35	0.73
Circular Orbit Search ($m = 0.2$)	0.35	0.49
Elliptical Orbit Search ($m = 0.2$)	0.19	0.33

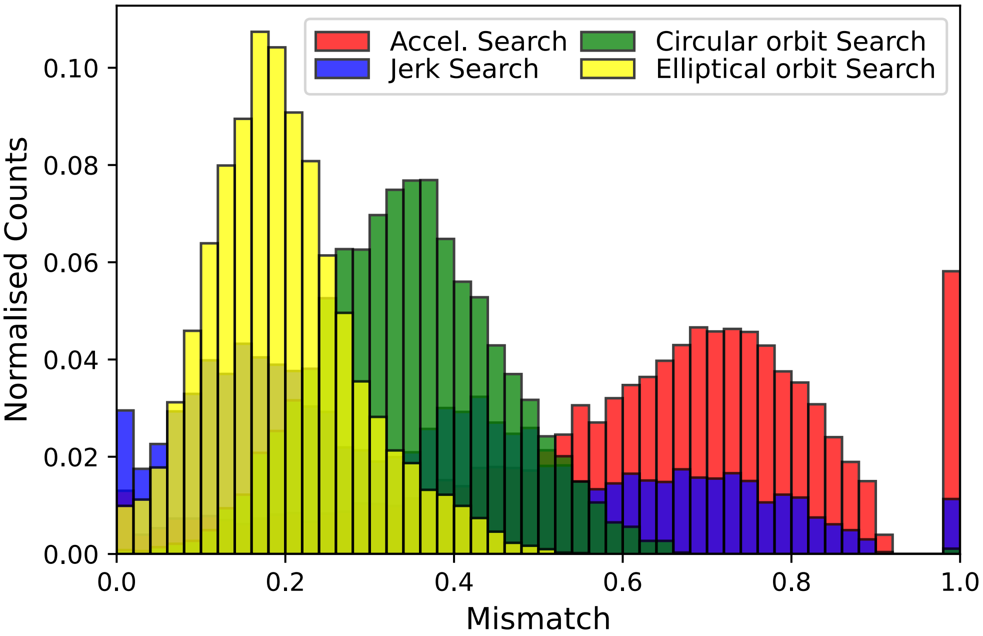


Figure 5.5: 1-D histogram of fractional signal to noise recovered in terms of mismatch values from a polynomial based acceleration and jerk search compared to Keplerian searches assuming a circular orbit binary and an elliptical orbit binary. These simulations were carried out on 72-minute observations of mildly recycled (15 ms) PSR-SBH binaries ($m_{\text{companion}} = 8\ M_{\odot}$) with an orbital period range of 6-12 hours in eccentric orbits (e=0.1).

DM $= 48.92\,\mathrm{pc\,cm^{-3}}$ to form a timeseries. We then downsampled the data by a factor of 16 to 1280 µs. We compared all the pipelines on two different integration lengths of 90-minute and 45-minute duration. Additionally, we also searched the data by shifting the starting point of the observation in steps of 100s. This effectively helped us sample different orbital phases of PSR J0737-3039 and test its detectability using different pipelines. For our searches using the PRESTO pipeline we used $z_{max} = 1200$ for acceleration search which using equation 5.33 corresponds to a maximum acceleration value of $1120.95\,\mathrm{m\,s^{-2}}$ for the fundamental in the 45-minute observation and $280.24\,\mathrm{m\,s^{-2}}$ in the 90-minute observation. Similarly, for jerk search we used a $z_{max} = 600$ and $w_{max} = 1800$ which using equation 5.34 corresponds to maximum acceleration of $560.47\,\mathrm{m\,s^{-2}}$ and jerk $0.62\,\mathrm{m\,s^{-3}}$ in the 45-minute observation. For the 90-minute observation, this corresponds to a maximum acceleration of $140.2\,\mathrm{m\,s^{-2}}$ and a maximum jerk of $0.078\,\mathrm{m\,s^{-3}}$. As mentioned in the previous section our jerk search range was limited by the amount of available RAM in the system. For the elliptical search template-bank, we set the maximum eccentricity to be equal to 0.08 and longitude of periastron was set to vary from 0-2π. For all the template-banks, the minimum pulsar mass and maximum companion mass was kept fixed at $1.4M_{\odot}$ and 1.6 $M_{\odot}$ respectively. We also searched over all possible values of initial orbital phase and inclination angles. We tweaked the coverage, mismatch, spin-period and orbital period values for different template banks in order to do a feasible search. These values are shown in table 5.4. Our results are shown in figure 5.6. In the top panel we present the FFT detection significance in the y-axis which is the sigma value reported from the 'ACCEL' files in PRESTO. The sigma reported is the probability that a given signal might be due to noise, but expressed in terms of equivalent Gaussian sigma. For the detections from the template-bank pipeline, we applied a time-domain resampling scheme (explained in section 5.5.1) on the dedispersed timeseries using the orbital template that gave us highest significance detection in the FFT. We then ran the code ACCELSEARCH from PRESTO with a $z_{max} = 2$ and compared the sigma value reported to those obtained from the standard acceleration and jerk search. The x-axis shows the starting point of our observation. Our results show that for 45-minute observations, acceleration search can partially recover the signal for some orbital phases but the detection is relatively weak for most of the orbital phase. As expected, jerk search improves the detection significance substantially. However, its sensitivity is still worse compared to full orbital searches. We don't notice a significant difference between the detection significance from elliptical and circular orbit searches. This is likely due to the fact that for detecting Double-Neutron Star systems for short observations in the $T_{\mathrm{obs}}/\mathrm{P}_{\mathrm{orb}} = 30\%$ range, a five-parameter search is not needed. For the 90-minute observation, we notice a significant difference in the performance of each pipeline. We notice the limitation posed by the assumption of constant acceleration and/or jerk in a long observation which hinders detection. Here, we also see the improved sensitivity provided by the elliptical search pipeline compared to a circular search. We would like to stress here some of the limitations of this comparison. The FFT candidate significance reported from the 'ACCEL' files includes a correction for the number of independent frequencies/fdots searched whereas the significance from the template-bank pipeline

Table 5.4: Orbital Parameters used to generate the template bank for tests on the observation of PSR J0737-3039A.

Input Parameter	$T_{\rm obs = 45min}$		$T_{\rm obs = 90min}$	
	Circular Search	Elliptical Search	Circular Search	Elliptical Search
Coverage (η)	0.9	0.9	0.9	0.5
Mismatch (m)	0.05	0.05	0.2	0.6
Max.Spin Period ($\rm P_{spin}$, ms)	20	22	22	22.5
Orbital Period range ($\rm P_{orb}$, hrs)	(2, 7.5)	(2.45-2.457)	(2.33-15)	(2.4-2.46)
No. of trials	65,489	52,981	615,800	3,00,000

was computed from a very small z search in the fourier domain after correcting for the orbit, therefore it does not include a correction for the number of trials. This would likely lead to an overestimation of the performance improvement provided by the full orbital searches. Another limitation is the range of acceleration and jerk search values used in the 90-minute observation. We expect an improvement in performance with a wider Acceleration and Jerk search range, however we still expect it to perform much worse than full orbital searches due to the varying jerk during an observation. In the bottom panel of figure 5.6, we report the folded significance calculated from the code PREPFOLD from PRESTO as a function of the starting point of the observation. For comparison, we also plot the significance estimated from folding the ephemeris of the pulsar. The significance reported here for all search pipelines assumes a single-trial search. As expected, the profiles of the folded significance is similar to the FFT significance values. The main advantage of the Keplerian search pipeline is that we need not rely on favorable orbital phases for detection. The slight non-uniform sensitivity of the template-bank for different orbital phases is due to our probabilistic coverage of the parameter space.

5.4 Suitability for targeted and blind observations

In this section, we estimate the feasibility of applying a coherent search for circular and elliptical orbit binaries in blind pulsar surveys like the High Time Resolution Universe (HTRU; Keith et al. 2010) as well as a targeted Globular Cluster (GC) observation of Terzan5.

5.4.1 HTRU-South Low Latitude Survey

High Time Resolution Universe South Low-Latitude Survey (HTRU-S Lowlat) is one part of the HTRU Survey that focuses on the inner Galactic plane covering Galactic longitude $-80^\circ < l < 30^\circ$ and Galactic latitude $|b| < 3.5^\circ$. The observations were recorded with an integration time of 72 minutes with a bandwidth of 400 MHz split into 1024 channels. The data was recorded using the 21-cm multi-beam receiver along with the Berkeley–Parkes–Swinburne Recorder (BPSR) backend at the 64-m Parkes Radio

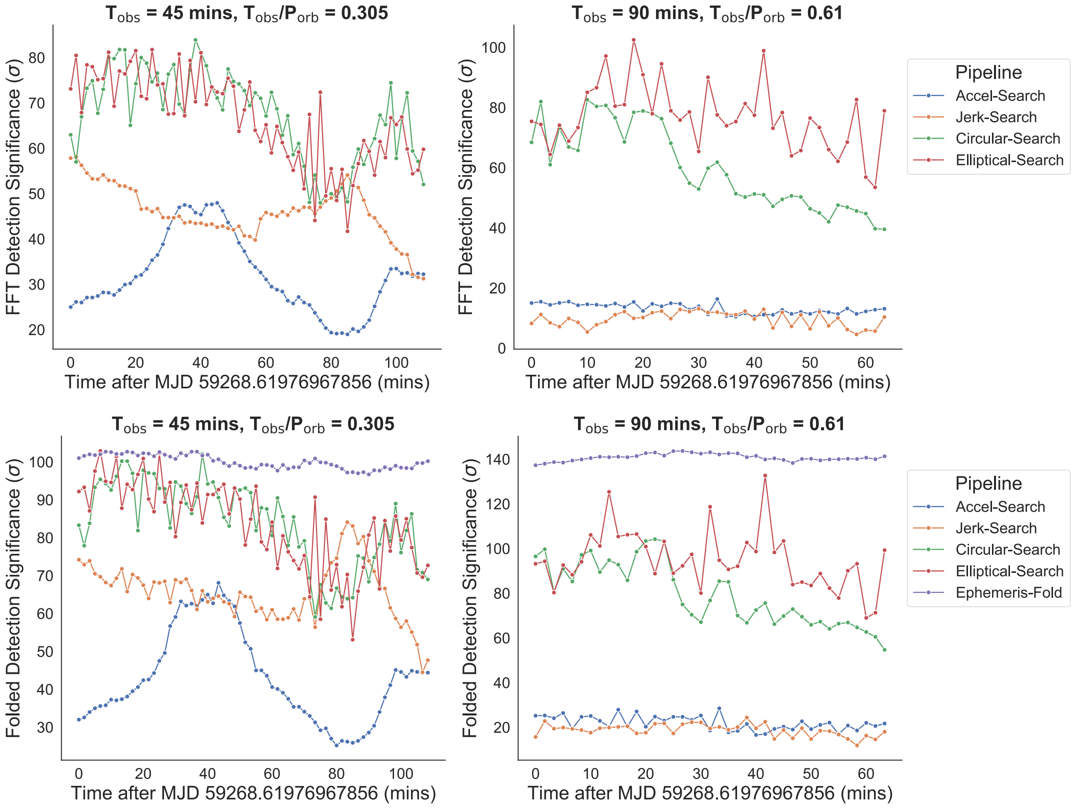


Figure 5.6: Results from running an acceleration, jerk, template bank circular and elliptical binary search pipeline on different integration lengths across a 2.6 hour Parkes observation of PSR J0737-3039A. Left panel shows results from searching over 45-minute observations and right panel shows results from searching over 90-minute observations. The top row shows the FFT detection significance calculated from ACCELSEARCH as a function of starting point of the observation incremented by 100 seconds. The bottom row shows folded detection significance of the pulsar calculated from PREPFOLD. For comparison, in the bottom panel we also show the detection significance obtained from folding the ephemeris of the pulsar in purple.

Telescope. We refer the reader to Keith et al. (2010) for a full list of the observational setup and system configuration. The main aim of this survey was to find low-luminosity pulsars and new relativistic binary pulsars that are expected to be found close to the Galactic disk. HTRU-S Lowlat has 1230 pointings, with each pointing consisting of 13 beams. More than 100 new pulsars have already been discovered in HTRU-S Lowlat using a segmented acceleration search pipeline. We refer the readers to Ng et al. (2015) and Cameron et al. (2020) for the initial list of discoveries and timing solutions. Given that we have already completed a first pass acceleration search over the entire dataset, this gives us a strong motivation to explore new techniques that can open up previously unexplored parameter spaces. We first start by estimating the computing time that would be required to search for non-recycled PSR-SBH binaries in HTRU.

5.4.2 Searching for Non-Recycled PSR-SBH Binaries in HTRU

In the standard binary stellar evolution model (Yungelson & Portegies Zwart, 1998; Voss & Tauris, 2003), we expect the black hole to form first, followed by the pulsar at a later epoch. Therefore, in this formation channel we expect to find a slow spinning non-recycled pulsar orbiting a black hole in a wide eccentric orbit. Recent population synthesis work by (Chattopadhyay et al., 2021) has shown that MeerKAT is expected to find 0-30 PSR-BH systems and SKA is expected to observe 1–60 PSR+BHs in the Galactic field with most of them expected to contained unrecycled normal pulsars. A similar earlier study from Shao & Li (2018) has also estimated that there might be ∼3-80 PSR-BH binaries in the Galactic disk with about 10 % that could be detected by Five-hundred-meter Aperture Spherical radio Telescope (FAST). The number of orbital templates is highly sensitive to the chosen spin frequency of the template bank. In the circular case, templates scale as f^3 and in the elliptical case, templates scale as f^5. Therefore, a natural place to apply the template-bank algorithm is for the search of normal unrecycled pulsars in compact binary systems. For all our calculations in this section, we fix the spin period in our template bank at $P_{\rm spin} = 100$ ms. HTRU observations are natively sampled at 64 µs, however given that we are targeting non-recycled pulsars with a spin-period greater than or equal to 100 ms, it is safe to downsample the data by a factor of 16 to 1024 µs. This gives us access to up-to 32 harmonics of the pulsar signal below the Nyquist limit. Each observation is then dedispersed to dispersion measures of 0 to 3000 $\mathrm{pc\,cm^{-3}}$ typically consisting of 1126 DM trials assuming a DM smearing tolerance of 10%. Using this and assuming we would fold at least 1000 candidates per beam, we estimate the total time it would take for our analysis in table 5.5. We calculate the cumulative time it would take to search using the random template-bank algorithm and folding on the full length HTRU observation as well as segmented searches on halves and quarter length observations. The total 'Search' time includes applying a channel mask, Fourier birdie zapping, dedispersion, Fourier de-reddening, FFT, harmonic summing up-to 16 harmonics and a peak detection algorithm. Total time in shorter observations includes searching over all the relevant segments (e.g. 2 and 4 segments for 36 and 18 minute observations respectively). Our search performance was benchmarked on an Nvidia P100 GPU. The next step is called folding, where once the data is dedispersed, we sum the data at a particular spin-period which is calculated based on the FFT period, orbital parameters and observation duration to form a pulsar candidate. We used the code PREPFOLD from PRESTO which also includes an optimization algorithm to find the best period and dispersion measure and is designed to be executed on a CPU. The folding performance was benchmarked on an Intel Xeon Gold 6140 processor. The total computation time reported in the table assumes the availability of 100 GPUs and 500 CPU cores. We added an overhead of 40% for the total folding time to account for data transfer times, I/O, and non-linear scaling during parallel execution of folding. We fix the upper limit for the orbital period in our calculations to be $P_{\rm orb} = 10T_{\rm obs}$ since an acceleration search would have already found binaries with orbital periods larger than this value. From these results, it is clear that for the vast majority of the parameter space, an elliptical orbit binary search for non-recycled pulsars is not feasible for a blind pulsar survey like HTRU. However, we notice that the regime between 2.5-10 $T_{\rm obs}$ could be feasible to search with a maximum eccentricity of 0.1. For example, an elliptical search for binaries between 1.5 to 6 hour orbital periods can be done in 145 days for the entire HTRU-S Lowlat survey searching across both half-length (36 minutes) observations. Any higher order eccentricity search is impractical with current and near-future computational resources. On a more optimistic note, Bagchi et al. (2013) showed that highly eccentric binaries ($e > 0.5$) are on average more favorable for detection compared to mildly eccentric systems ($e = 0.1$) as the pulsar spends most of its time away from periastron and has a low line-of-sight acceleration for most of its orbit. Therefore, from both a physical and computational point of view, it is more favorable to stick to maximum eccentricity of 0.1. Additionally, from table 5.5 we notice that a circular binary search is computationally feasible and can open up interesting parameter spaces that are yet to be investigated. We notice that the searches between 1.7-10 $T_{\rm obs}$ can be analysed with current resources and even non-detections can be used to calculate limits on the existence of relativistic binary systems within our Galaxy. For searches longer than 50 days, it may be desirable to reduce the total number of orbital templates (usually by a factor of 2) by applying the stochastic template-bank algorithm. This requires the investment of a few days to 10 days worth of prior computing time using multiple cores on a HPC cluster in order to prune overlapping orbital templates. An implementation of this algorithm can also be found in our software repository.

5.4.3 Searching for Millisecond PSR-SBH Binaries in HTRU

The formation of a millisecond pulsar (MSP) orbiting a black hole is expected to be rare. One possible scenario is the so-called reversal mechanism where under certain conditions the pulsar is formed first and is later spun-up by accretion during the red giant phase of the secondary star (Sipior et al., 2004; Pfahl et al., 2005). Discovering such systems more desirable as MSPs tend to be more precise timers than slow ‘normal’ pulsars (Verbiest et al., 2009). Another possibility is the dynamical formation scenario due to exchange interactions. This is more likely in regions of high stellar density like the Galactic centre. While interstellar scattering is expected to hamper detections of

Table 5.5: Total Computing time required to search for non-recycled PSR-SBH binaries ($P_{\text{spin,min}} = 100$ ms.) using a coherent circular binary and elliptical binary search in the entire HTRU-S Lowlat Survey. Total 'Search' time includes applying a channel mask, Fourier birdie zapping, dedispersion, Fourier de-reddening, FFT, harmonic summing up-to 16 harmonics and a peak detection algorithm. Our searches were benchmarked on a Nvidia P100 GPU and the folding was benchmarked on an Intel Xeon Gold 6140 processor using PREPFOLD from PRESTO .

t_{obs}	Min. P_{b}	Max. P_{b}	Mass$_{\text{companion}}$	Max. Eccentricity[a]	Circular Binary Search			Elliptical Binary Search		
(mins)	(h)	(h)	($M_\odot$)		Search(d)	Folding(d)	Total (days)	Search(d)	Folding(d)	Total (years)
72	3	12	8	0.1	16.2	35.4	51.6	1613.6	35.4	4.5
72	2	12	8	0.1	229.6	35.4	265	660,079.5	35.4	1808
36	1.5	6	8	0.1	5	19	24	145	19	0.45
36	1.0	6	8	0.1	68.7	19	87.7	59,717.8	19	163.7
18	0.42	3	8	0.1	70.8	10	80.8	108,780.9	10	298

[a] Applicable only for Elliptical Orbit Searches.

PSR-BH binaries near Sgr A* at L-band in HTRU-S Lowlat, it is still important to search for such systems due to their potential scientific impact. Given that elliptical binary searches were already unfeasible for searching for unrecycled PSR-SBH binaries, the situation is expected to be worse for recycled PSR-SBH binaries. However, circular binary searches are still feasible and we provide a few examples of search-setups in table 5.6. We focus mainly on the spin-period regime between 13-20 ms. MSPs rotating faster than this are still computationally unfeasible to search for PSR-SBH binaries. An alternate possibility is to reduce the companion mass limit and focus searches for double neutron-star and pulsar-white dwarf binaries. This would help in reducing the spin-period limit and search for more recycled pulsars. Knispel (2011); Allen et al. (2013) and Knispel et al. (2013) used this set-up to search for pulsars in PALFA (Cordes et al., 2006) and PMPS (Manchester et al., 2001)) respectively using the volunteer distributed computing resources available from the Einstein@Home project. Our calculations indicate that the regime between 5-10 T_{obs} is feasible to search for MSP-SBH binaries in HTRU with spin-periods greater than 13 ms using a circular binary search. The entire analysis would take anywhere between 6-18 months depending on the chosen search setup. Additionally, having a higher spin-period threshold at 20 ms could help us investigate binaries with orbital period in the 3-10 T_{obs} regime. Our benchmarks were done assuming the same computational resources as in the previous section. In this case, we downsampled the data to 256 µs which gives us access to at least 16 harmonics of a hypothetical 13 ms pulsar signal. The total search time includes 1876 dedispersion trials assuming a DM smearing tolerance of 10% from 0 to 3000 $\text{pc}\,\text{cm}^{-3}$.

5.4.4 Searching for MSP-SBH binaries in Globular Clusters

Globular Clusters (GCs) are one of the ideal locations to apply the template-bank algorithm. GCs produce orders of magnitude more MSPs than the Galactic disk and since most MSPs are expected to be binaries, we expect to find several compact bi-

Table 5.6: Total Computing time required to search for MSP PSR-SBH binaries using a coherent circular binary search in the entire HTRU-S Lowlat Survey. For searches longer than 50 days, we recommend applying the stochastic template-bank algorithm and pruning overlapping templates. This usually reduces the number of templates by a factor of 2 at the cost of an extra week of up-front computing time.

t_{obs}	Min. P_{b}	Max. P_{b}	$\mathrm{Mass}_{\mathrm{companion}}$	$\mathrm{P}_{\mathrm{spin}}$	Circular Binary Search		
(mins)	(h)	(h)	($M_{\odot}$)	(ms)	Search(d)	Folding(d)	Total (days)
72	6	12	8	13	355.6	214.0	569.6
72	6	12	8	20	99.0	214.0	313.0
36	3	6	8	20	99.0	86.5	185.5
36	2	6	8	20	400.3	86.5	486.8
18	1	3	8	20	120.0	67.3	187.3
9	0.42	1.5	8	20	99.3	58.0	157.3

nary pulsars in GCs. There are currently 223 known pulsars in 36 globular clusters[4] Additionally, the advantage of processing GCs is that the DM of the cluster maybe well known which means most of the compute power can be invested in exploring the binary parameter phase space. Any new candidate can also be confirmed potentially using archival observations of these clusters. Therefore, we could have a timing solution that spans across several years within a short period of time after detection, thus enabling us to get scientific results quicker. In the context of PSR-SBH binaries, analogous to the Galactic centre, the high stellar density in GCs opens up the possibility of three-body interactions where a neutron star can gain a companion by exchanging with a primordial binary and subsequently be spun up to become an MSP (Hills, 1976; Sigurdsson & Phinney, 1995; Ivanova et al., 2008; Clausen et al., 2014). Another possibility is the formation a MSP in orbit with an intermediate-mass black hole (IMBH: $M \sim 10^2$ - $10^4 M_{\odot}$). The encounters between MSPs and IMBHs could result in the MSP being significantly displaced from the core (Colpi et al., 2003) to form a MSP-IMBH binary (Devecchi et al., 2007). We investigate the compute time required to search for such binaries using both a circular and elliptical orbit search in table 5.7. We used a 7.2 hour archival observation of Terzan5 taken at the 100-m Robert C. Byrd Green Bank Telescope (GBT) for our benchmarks (Prager et al., 2017). The observation was recorded at a central frequency of 2000 MHz (S-band) using a bandwidth of 800 MHz split into 128 channels. The data recorded by the Green Bank Ultimate Pulsar Processing Instrument (GUPPI) were full-Stokes with a sampling time of 10.24 µs and 512 channels, each coherently dedispersed to a DM of 238 pc cm^{-3}. For our benchmarks, we downsampled the data to 81.9 µs and assumed 50 dedispersion trials and the same computational resources as the previous section. For the elliptical searches, we restrict the maximum eccentricity to 0.1. Any higher order eccentricity search is still not

---

[4] Refer to `http://www.naic.edu/~pfreire/GCpsr.html` for an updated list of the catalog.

Table 5.7: Total Computing time required to search for PSR-SBH binaries using a coherent circular binary and elliptical binary search in a GBT observation of Terzan5. Duration includes time taken for searching and folding. These numbers assume 100 Nvidia P100 GPUs and 500 CPU cores. The total time taken for observation shorter than 7.2 hours includes time taken to cover all search segments.

t_{obs}	Min. P_b	Max. P_b	Circular Binary Search			Elliptical Binary Search			
(hrs)	(h)	(h)	P_{spin} (ms)	$\text{Mass}_{\text{companion}}$ ($M_\odot$)	Duration (days)[a]	P_{spin} (ms)	$\text{Mass}_{\text{companion}}$ ($M_\odot$)	Max. Eccentricity[a]	Duration (days)[a]
7.2	36	72	10	10	1.3	40	10	0.1	27.1
3.6	18	36	10	10	0.24	40	10	0.1	0.95
1.8	9	18	10	10	0.06	40	100	0.1	2.2
1.0	5	10	10	1000	2.0	20	10	0.1	0.2
0.5	2.5	5	1	1000	0.14	15	10	0.1	0.09
0.5	1.5	5	10	100	1.69	20	10	0.1	2.56
0.5	1.0	5	5	1.6	0.9	40	1.6	0.1	0.9

[a] Includes time taken for search and folding.

computationally feasible. As expected processing longer integration times are computationally more challenging. Based, on our benchmarks a circular-orbit search between 5-10 T_{obs}, i.e. 36-72 hour orbital period binaries for a 7.2-hour observation up to a companion mass of 10 $M_\odot$, would take about 1.3 days assuming 100 GPUs for searching and 500 CPU cores for folding. We also provide rough estimates for segmented searches. The total processing time for observations shorter than 7.2 hours includes searching all search segments (for e.g. 2 segments for a 3.6-hour observation). Elliptical binary searches are not feasible for very long integration times. In order to make the elliptical binary search feasible, we need to either search on shorter length observations or have a higher threshold for the spin-period in our template-bank. We have given a few examples of possible parameter-space combinations that can be explored with current computational resources. This is given in the right hand side of table 5.7. Some examples readers might find interesting include an elliptical binary search on a 1.8 hour observation with a spin-period threshold of 40 ms and a maximum companion mass of 100 $M_\odot$. Another example would be a circular binary search between orbital period of 5-10 hours with a maximum companion mass of 1000 $M_\odot$. Yet another possibility to reduce computation time is to focus on double neutron star searches. This would help to investigate faster spinning pulsars in more compact orbits. For example, using a circular binary search for a 30 minute observation, we could search all 14 segments (total length = 7.0 hrs), with an orbital period range of 1-5 hours (2-10 T_{obs}) with a maximum companion mass of 1.6 $M_\odot$ in a day.

5.5 Software and Implementation Details

Our search code has been designed to run on Nvidia-GPUs. It is built on top of the GPU-accelerated PEASOUP[5] pipeline that does an acceleration search (Johnston & Kulkarni, 1991). The main pipeline also implements dedispersion through the

---

[5] `https://github.com/ewanbarr/peasoup/`

DEDISP library (Barsdell et al., 2012), Fourier de-reddenning, FFT, incoherent harmonic summing and a peak detection algorithm for candidate selection. Our version of the software called 5D-PEASOUP[6] replaces the acceleration search with a coherent search across all Keplerian parameters and additionally runs a candidate reduction algorithm by grouping together multiple detections of the same candidate.

5.5.1 Time-Domain Resampling Algorithm

We apply the nearest-neighbour resampling algorithm which is the standard resampling technique applied in most time-domain pulsar search pipelines including SIGPROC, PEASOUP and the Einsten@Home Radio Pulsar-Search pipeline (Knispel, 2011; Allen et al., 2013) . The difference here is that our Roemer delay equation contains all five Keplerian parameters. Our goal is to transform our reference frame from the observed timeseries at the telescope to the binary system's barycenter. After correcting for the dispersion delay caused by free electrons in the Inter-Stellar Medium (ISM) using the software DEDISP, we additionally need to account for the Roemer delay caused due to the motion of the pulsar around its companion (Blandford & Teukolsky, 1976). For a Keplerian orbit with non-zero eccentricity, this can be written as:

$$\Delta_R = \tau[(\cos E - e)\sin\psi + \sqrt{1-e^2}\sin E\cos\psi]. \tag{5.35}$$

As mentioned earlier calculating the sine and cosine of eccentric anomaly involves solving Kepler's equation iteratively. The alternative is to expand $\cos E$ and $\sin E$ as a power series shown in equation 5.15, 5.16. The order to which the power series should be expanded depends on the eccentricity range and the companion mass we want to be sensitive to. These models are referred to as ELL1 (Wex 1998, unpublished contribution to TEMPO, see also Lange et al. 2001), ELL2 and so on based on the number of retained terms in the power series. For the calculation of the template-bank, we used the ELL7 model which covers up to $e \approx 0.8$. However, in our resampling algorithm we implemented the BT model. Therefore, the time t' at the barycenter can be calculated by solving

$$t' = t - \tau[(\cos E - e)\sin\psi + \sqrt{1-e^2}\sin E cos\psi], \tag{5.36}$$

where we have ignored constant phase shifts caused by initial signal phase and the constant classical light travel time between the pulsar and the barycenter. The signal S is recorded at the telescope in discrete time steps $S = i \times t_{\text{samp}}$ where i is the bin number and t_{samp} is the sampling time and our goal is to calculate $S' = j \times t_{\text{samp}}$ which is the signal at the barycenter. We first start by calculating a zero offset ($t = 0$) for each orbital template, in order to shift the scale of the observed timeseries to match with the barycentered timeseries. This number is then subtracted from the Roemer delay, which is usually a rational number and then we calculate the nearest integer to this number and map the value in that bin to our resampled timeseries. A brief description of this can be found in algorithm 1.

---

[6] `https://github.com/vishnubk/5D_Peasoup`

Algorithm 1: Nearest-Neighbour Demodulation Algorithm

```
for K orbital templates in template bank do
    Calculate Mean Anomaly $M$ at time $t = 0$
        $M = \alpha$
    Solve for Eccentric Anomaly $E$ iteratively
        $E - e \sin E = M,$
    Compute Zero-Offset $t_0$
        $\mathrm{t}_0 = \frac{\tau[(\cos E - e)\sin\psi + \sqrt{1-e^2}\sin E \cos\psi]}{t_{\mathrm{samp}}}$
    for $i$th bin in observed timeseries do
        $M = \Omega((i \times t_{\mathrm{samp}}) + \alpha)$
        Solve for $E$ iteratively
            $E - e \sin E = M,$
        $S'_j = S\left[j - \frac{\tau[(\cos E - e)\sin\psi + \sqrt{1-e^2}\sin E \cos\psi]}{t_{\mathrm{samp}}} - \mathrm{t}_0\right]_{\mathrm{int}}$
    end
    return $S'$ to calculate FFT
end
```

5.6 Discussions and Conclusions

A question not addressed in our work is the special case of targeted searches i.e. when one or multiple Keplerian parameters are known accurately enough that a search across that dimension is not needed. We refer the readers to section 3-B of (Dhurandhar & Vecchio, 2001) for a more detailed formalism of creating a template-bank in this scenario. This can significantly reduce the number of orbital templates required for searching. We expect this to be incorporated in a future version of the software. Another promising approach that requires further investigation in the context of radio pulsar searching is hierarchical searches. Nieder et al. (2020) described a multi-stage search where in the first stage the parameter space is investigated at a lower sensitivity (high mismatch), followed by searching in smaller regions of the parameter space around promising candidates with higher sensitivity. This technique was successfully used to discover a binary Gamma-Ray Pulsar in Fermi Large Area Telescope (LAT) data (Nieder et al., 2020). In this paper, we described the implementation for creating a template-bank to do a fully coherent search across three and five Keplerian parameters assuming circular and elliptical orbits, respectively. We demonstrated the extra sensitivity gained by applying template-bank pipeline compared to acceleration and jerk searches from PRESTO for simulated Pulsar stellar-mass Black hole binaries and the double pulsar PSR J0737-3039A. Trade-offs of this algorithm include a significantly longer computation time and reduced sensitivity to longer orbital period binaries due to the increased number of trials and isolated fast-spinning pulsars due to our high spin-period threshold. We also benchmarked the amount of time it would take to search for non-recycled PSR-SBH binaries in HTRU. We note that the regime between 2.5-10 T_{obs} should be feasible to search in 36-minute half length observations of HTRU-S Lowlat with a maximum eccentricity of 0.1 and a spin-period threshold of 100 ms. Searching across all five Keplerian parameters for millisecond PSR-SBH binaries on a survey like HTRU-S Lowlat is still not computationally feasible. However, circular binary searches could open up interesting parameter spaces here that are yet to be explored by acceleration and jerk searches. For example, assuming a maximum companion mass of 8 $M_{\odot}$, the regime between 5-10 T_{obs} should be feasible to search in full length HTRU-S Lowlat observations with a spin-period greater than or equal to 13 ms. By reducing the spin-period threshold to 20 ms, we could explore binaries with orbital period in the 3-10 T_{obs} regime. We expect the most applicability of the template-bank pipeline on targeted observations like globular clusters. Here, our benchmarks indicate that searching for mildly recycled pulsars orbiting an intermediate mass black hole is feasible for observations shorter than 2 hours with a maximum eccentricity of 0.1 in the template-bank. Additionally, we also open-source our circular and elliptical binary search GPU pipelines.

Acknowledgements

We would like to thank Norbert Wex and Benjamin Knispel for useful discussions. We would also like to thank Scott Ransom for providing us with the GBT observation of

Terzan5. The data analysis were performed on the OzSTAR national supercomputing facilities at Swinburne University of Technology and the HERCULES computing cluster operated by the Max Planck Computing & Data Facility (MPCDF). OzSTAR is funded under Astronomy National Collaborative Research Infrastructure Strategy (NCRIS) Program via Astronomy Australia Ltd (AAL). We would like to thank members of the open-source community for maintaining software packages that were directly used for our work including NumPy (Dewdney et al., 2009), Matplotlib (Hunter, 2007), emcee (Foreman-Mackey et al., 2013), corner (Foreman-Mackey, 2016), SymPy (Meurer et al., 2017) and CUDA (NVIDIA et al., 2020).

Chapter 6

Conclusion and future work

6.1 Summaries and Conclusion

6.1.1 HTRU-S Lowlat Survey

The primary aim of this PhD thesis was to reprocess and analyse observations of HTRU-S LowLat survey using new binary pulsar search techniques in order to maximise our chances for a PSR-BH binary detection. The main goals of this survey was to discover low-luminosity pulsars which was made possible due to the high time and frequency resolution of the survey and its long integration length of 72-minutes. An important secondary goal of the survey was to find short relativistic orbit pulsars close to the Galactic plane. Given that when this PhD started, a first-pass segmented acceleration search (Ng et al., 2015; Cameron et al., 2020) was already used to reprocess 94.1% of the survey, a considerable amount of this thesis is devoted to the development of new binary pulsar search techniques that can probe binary phase-spaces in LOWLAT observations which are inaccessible for a traditional acceleration search. With this goal in mind, we developed a new CPU and GPU-based pipeline called 3D-PEASOUP[1] which uses the template-bank algorithm to search coherently for circular orbit binaries. This work was built on previous works in gravitational-wave searches and radio pulsar searches in the Einstein@Home project (Knispel, 2011; Allen et al., 2013; Knispel et al., 2013). Using this new tool, we performed two pulsar searches on the full-length 72 minute observations of LOWLAT. One of these were aimed at finding Recycled PSR-BH binaries with orbital periods between 6-12 hours and the second search was aimed at finding Non-Recycled PSR-BH binaries with orbital periods between 3-12 hours. We have used this to reprocess 87.1% of the LOWLAT survey, discovering 20 new pulsars including a new MSP J1743−24 which is a rare intermediate spin-period pulsar in a 70.7 day orbit alongside a low-mass companion star or less likely a He-WD star. We also presented updated timing solutions for the binary PSR J1753−2819, a pulsar similar to PSR J1743−24 but in a much shorter orbit of 9.3 hours. We also discussed the uniqueness of the combination of spin-period and companion mass for both these pulsars by comparing them to other known binaries. Both these pulsars could likely be part of a separate population which orbit low-mass companion stars with intermediate spin periods. In addition, we also presented discovery parameters of 50 new pulsars which were independently found during our reprocessing that was missed by the first-pass pipeline (Ng et al., 2015; Cameron et al., 2020). A further 60 known pulsars were also detected which were previously missed. Based on our new detections, we also

---

[1] https://github.com/vishnubk/3D_peasoup

carried out a renewed evaluation of the survey yield. Our discoveries and redetections have now bridged the gap between the number of expected discoveries from simulations and observations. Cameron et al. (2020) reported a deficit of 25% of normal pulsar detections in LOWLAT which has now been reduced to a 9 % deficit. The number of MSP detections in the survey has been approximately doubled and is now consistent with the lower limit predicted from PSRPOPPY. We discussed in detail why these predictions should only be treated as rough estimates due to our lack of constrains in the luminosity distribution of normal pulsars especially at the lower end where fewer pulsars have been detected. Predictions for MSP population are similarly biased by the limited sample of Galactic-field MSPs especially at the time during which these simulations were first conducted. We believe it is likely that more new discoveries and redetections of known pulsars can be obtained by taking advantage of the full time resolution of LOWLAT and processing at a sampling rate of 64 µs. Based on our discovery and redetection rate, we expect LOWLAT to discover an additional 14 pulsars in the remaining part of the survey which is yet to be reprocessed. However, this number could vary as the part of the sky surveyed by LOWLAT has overlaps with the other ongoing Galactic plane-surveys like MGPS-L pulsar survey using MEERKAT (Padmanabh et al., prep) and the ongoing FAST Galactic Plane Pulsar Snapshot survey (GPPS; Han et al. 2021).

Given the extremely high number of discoveries from the reprocessing of the survey within a short period of time, a timing followup has not been possible yet for most pulsars except MSPs or confirmed binary pulsars. A dedicated 1-year timing campaign for these pulsars will be crucial to exploit their full scientific potential and these will be the immediate short term goals of this project along with the reprocessing of the remaining portions of the survey. A high priority is currently being placed for observations of PSR J1743 − 24 which is currently being followed up by both Parkes and Effelsberg Radio telescopes, and PSR J1142 − 62, with the former confirmed to be a binary and will now be observed in timing mode and the latter showing indications of being part of a binary orbit. As explained in section 3.3.6, the population of pulsars found in the reprocessing of LOWLAT on average have larger detected DM, therefore its likely that we are probing pulsars located further away from us. Discoveries from surveys like LOWLAT due to its uniform sensitivity and wide coverage are crucial in order to get an unbiased census of the Galactic-plane pulsar population. This will be useful in order to plan future pulsar survey strategies with the SKA.

A major achievement of this work is the increased sensitivity we now have for relativistic Pulsar-BH Binaries. To the best of our knowledge, this is the most comprehensive PSR-BH binary search done in a Galactic plane survey. We have managed to improve our sensitivity towards such systems by a factor of **2-2.5** which was achieved entirely due to a change in the algorithm. However, despite this, we report no detections from a PSR-BH binary yet and no new relativistic pulsar discoveries. We are confident that the pipeline performs as expected given our extensive number of tests on simulated and real observations as well as higher SNR redetections of known relativistic binary pulsars like PSR J1757−1854 (Cameron et al., 2018) and PSR J1756−2251 (Faulkner et al., 2005).

We have used our non-detections to place limits on short orbital period PSR-Stellar

Mass BH binaries and we also comment on the detectable fraction of PSR-BH binaries from our searches. Our non-detections indicate that the existence of nearby PSR-BH binaries ($d \leq$ 1kpc) with circular orbits and orbital period 4-24 hours is highly unlikely. The possibility of PSR-BH binaries having significantly eccentric orbits or circular orbits shorter than 4 hours cannot currently be ruled out due to them being outside our search-range. We leave these parameter spaces to be explored by future reprocessing efforts.

A major drawback of our search pipeline is the requirement of more computation trials due to the additional binary search parameters. Since our goal was to be sensitive towards PSR-BH binaries, due to computational reasons, we were forced to have a very high threshold for spin-period of 13 ms in our template-bank. This has the unfortunate drawback of higher mismatch towards isolated and binary MSPs rotating faster than 13 ms. Additionally, our increased binary trials limited us to analyse decimated observations of LOWLAT and not to exploit the full sensitivity of the survey. This can partially be addressed by future efforts which will be aided by advances in computation especially GPU technologies. Processing pulsar surveys is a computationally demanding task and given the extensive speed-up that can be achieved by doing searches on GPUs (see section 3.2.10), we believe that this will likely be the preferred processing option in future pulsar surveys. Our experience shows that current processing bottlenecks in our search pipelines is due to the time required for folding pulsar candidates which are still done on a CPU. Therefore, more work needs to be in done in this field and to study if folding can be similarly accelerated by running on a GPU.

6.1.2 Pulsar Candidate Identification using SGAN

The output of a pulsar search pipeline is a pulsar candidate-a possible detection of a new pulsar. Modern pulsar surveys like HTRU-S Lowlat typically produce around 40 million pulsar candidates in one processing run, thereby making manual inspection infeasible. We have elaborated more on the reasons for this in chapter 4 of this thesis. ML techniques are a promising solution to this problem. One of the challenges in building a pulsar ML classifier is the so called class-imbalance problem i.e out of the total 40 million pulsar candidates, less than 1% are expected to be real pulsar detections (including harmonics and duplicate detections of the same pulsar). Supervised ML classifiers require that both datasets be similarly balanced. The traditional approach of solving this problem has been to either under-sample the number of radio inteference candidates or over-sample the number of pulsar candidates using techniques like GANs (Guo et al., 2019), or Synthetic Minority Oversampling Technique (SMOTE) (Bethapudi & Desai, 2018). In chapter 4, we presented a novel semi-supervised machine learning pipeline called SGAN that can learn from both labelled and unlabelled data. The latter is crucial since this saves a lot of human time required to manually label pulsar candidates. Our results are a major improvement over current approaches when few labelled data is available and this is a promising solution for early stages of future pulsar surveys. Our work is the first implementation of a Semi-Supervised ML classifier in the literature for pulsar candidate classification. Future pulsar surveys using MeerKAT and SKA will likely produce more pulsar candidates, thereby ML classifiers based on semi-supervised learning will likely gain further precedence with time as maintaining a well labelled candidate dataset becomes increasingly challenging. Some of the challenges of semi-supervised learning include the higher training time required compared to past approaches. Our pipelines focuses more on final performance metrics rather than speed of classification. An alternative approach that focuses more on the latter is a Gaussian Hellinger Very Fast Decision Tree (GH-VFDT, Lyon et al. 2016). We believe that depending on the future needs of pulsar surveys, both these approaches will likely have its own unique applications.

6.1.3 Coherent Search for Relativistic Binary Pulsars across all five-Keplerian Parameters- The ultimate brute-force pipeline

In Chapter 5, we pushed the envelope of sensitivity regained from binary pulsar search pipelines further by presenting a new GPU-based pulsar search pipeline that can search coherently across all five-Keplerian parameters. This is the first implementation of a fully coherent search for elliptical orbits in literature. We demonstrated the extra sensitivity regained from this pipeline by comparing it an acceleration, jerk search and the template-bank pipeline assuming a circular orbit for observations of PSR J0737−3039A and simulated observations of Pulsar-Stellar Mass Black hole binaries. Approaches like these are crucial to discover relativistic binary pulsars in long observations where polynomial searches tend to lose sensitivity. The main challenge of this approach will be its computational feasibility. Our benchmarks indicate that this pipeline will be useful for targeted binary pulsar searches in Globular Clusters (GC) where the DM of the cluster is well known or targeted observations of the Galactic centre where we expect to find a lot of binaries due to the high stellar densities in this environment which aid exchange interactions. A feasible yet promising application of this pipeline would be in the search of pulsars orbiting intermediate mass black holes in the 5-10 T_{obs} regime in GC data. Another promising future application of this pipeline will be in the search of pulsar companions around known white-dwarf binaries (for e.g. Athanasiadis et al. 2021) or the follow-up of known eccentric binary pulsars close to periastron where standard assumption of constant acceleration or jerk breaks down. Given the current and near future growth of computation, its unlikely that this approach will be feasible for blind pulsar surveys like LOWLAT but we believe this pipeline will be a complementary and unique tool to standard pulsar search pipelines for targeted searches. One of the additional advantages of using the template-bank pipeline is that it can place limits on the existence of population of pulsars with certain orbital period, and companion mass ranges since Keplerian parameters are included in our search range. This information will likely provide strong observational constrains in the future and will be useful in order to enhance our understanding of binary stellar evolution.

6.2 Future Work

In terms of radio pulsar astronomy, we are currently living through an exciting time period with the advent of new telescope facilities like MeerKAT, FAST and eventually SKA all of which will undoubtedly increase the number of known pulsars and produce exciting scientific results. The gain in sensitivity due to bigger telescopes and better electronics will likely be the determining factor for new discoveries in the near future. This thesis however takes a different approach and investigates better algorithms and techniques that can probe the population of pulsars that are otherwise hard to detect. Such approaches will be crucial in order to minimise our selection effect towards relativistic binaries and once discoveries from standard approaches have dried up. Below, we give some details of some of the near-term goals of the projects undertaken during this PhD and some possible future research paths that can be undertaken based on our results.

6.2.1 Continued reprocessing of HTRU-S Lowlat

We have currently processed 87.1% of the LOWLAT survey with the template-bank pipeline. Our near-term goal would be to complete the processing of this survey which would then help us to evaluate the total survey yield of LOWLAT. Based on our non-detections, we are also currently working on improving our limits for the population of PSR-BH binaries using results from theoretical modelling of PSR-BH binaries in the galaxy (Shao & Li, 2018; Chattopadhyay et al., 2021). Results from these along with any additional discoveries from the remaining portion of the survey will be the subject of a future publication. Looking ahead, there are still many regions of the binary phase space that are yet to be probed in LOWLAT data. Some examples include a segmented template-bank search using the full time resolution of the data which from our past experience has yielded a significantly higher number of pulsar detections compared to searches on decimated data. Another possibility would be a search specifically designed for finding either DNS systems or PSR-WD binaries which will help us search for pulsars rotating faster than 13 ms.[2] Yet another possibility is to search for ultra compact binaries with orbits in the range of minutes in segmented observations of LOWLAT. These orbits are expected to be highly circular and will be greatly aided by our circular orbit search pipeline. With our ability to directly search over Keplerian parameters, and due to that fact these searches can be done on GPUs, a rich and diverse set of searches can be done in the near future on radio pulsar observations.

6.2.2 Follow-up of discoveries from our work

Yet another near term goal of this project is to complete the 1-year initial timing of all the pulsars presented in this work. This will be important in order to calculate physical parameters like spin-down rate which will give us the characteristic age and magnetic field strength of the pulsar. Regular monitoring of this pulsars will also help

---

[2] This was limit for the fastest spinning pulsar in our searches.

us detect any possible long-term emission instabilities of these pulsars like nulling, intermittentency or mode changing. The integrated pulse profile of PSR J1743−24 is quite complex with multiple components. Therefore, this warrants a more detailed investigation using polarisation data which would help us better understand its magnetic configuration and emission geometry. PSR J1609−5225 has the highest detected DM of 993(10) $\mathrm{pc\,cm^{-3}}$ in our discovery sample and its integrated pulse profile shows a strong scattering tail. We presented results from modelling this and measured a scattering timescale of $\tau = 77 \pm 25\,\mathrm{ms}$. Observations at high frequencies and wider bandwidth for this pulsar using UWL will be useful in order to minimise the effect of interstellar scattering and to calculate its scattering spectral index. Besides this, we will also be regularly monitoring our ClassA candidates described in section 3.3.5, two of which showed a high degree of acceleration and could therefore be binary pulsars which are either eclipsing or difficult to detect due to the impact of RFI. Confirmed pulsars north of -30° once they have been gridded at Parkes will be timed by the combination of Jodrell Bank, Sardinia and Effelsberg Radio telescopes in the future.

6.2.3 Improvements in Binary Pulsar- Search Algorithms

The binary pulsar search algorithms developed and implemented in this PhD still have rooms for improvement. A promising approach would be to implement hierarchical template-bank searches where initially we search for pulsar signals with a high mismatch and follow-up on promising signals that survive this stage with a template-bank that has a finer mismatch. An implementation of this approach was recently done by Nieder et al. (2020) which was used to search for Binary Gamma-Ray Pulsars. Hierarchical template-bank searches have a lower computational load compared to the fully coherent brute force search and it can also open up a phase-space that hasn't yet been explored in radio data. Another promising approach is the "pruning-method" first proposed by Zackay (2017). This algorithm processes data in smaller time segments and only retains candidates that increase in power by $\sqrt{2}$ with each addition of adjacent time segments. This can in principle help filter out several unwanted noise/interference candidates. A blind test comparing the pruning method to the template-bank algorithm on real observations will be of great interest and is a promising future research project.

Appendix A

Redetections of 60 known pulsars missing in the first pass processing.

Table A.1: Catalog listing 60 unique known pulsars that were missed during the initial processing of LOWLAT. We list here the topocentric spin-period, DM, the orbital template, and the corresponding htru pointing and beam where the detection was made.

psrname	pointing/beam	period_topo	dm_topo	sigma	porb	asini	t0
J1020−6026	2012-07-18-04:00:54/03	140.4877	448.06	13.1	25483.374940948	0	56125.916670482977
J1031−6117	2012-07-18-05:13:39/07	306.4130	504.26	11.0	23484.8050807693	0.13974750041962	56125.954865347172
J1052−5954	2011-05-05-05:55:13/10	180.5951	511.79	7.4	23484.8050807693	0.13974750041962	55685.983731087916
J1107−5947	2011-01-10-14:27:09/13	1516.8204	107.29	30.9	33308.7019947862	1.40841090679169	55571.438239627081
J1128−6219	2011-07-29-00:04:30/01	516.3971	694.38	8.9	26300.138216697	3.28635787963867	55770.880104102755
J1216−6223	2011-12-11-16:22:18/06	374.1119	797.52	8.4	33263.0986177139	1.06929671764374	55906.507634006994
J1233−6312	2011-01-03-15:12:08/04	564.7719	389.04	15.0	22749.5515469829	0.27001881599426	55564.485566675598
J1233−6344	2012-04-01-10:19:04/02	756.7896	500.72	10.7	24839.418691522	0.49663379788399	56018.178475103137
J1248−6344	2011-04-25-08:37:51/02	198.3301	435.51	13.0	23484.8050807693	0.13974750041962	55676.096670902727
J1301−6310	2011-12-22-15:34:56/09	663.7576	87.19	13.1	24839.418691522	0.49663379788399	55917.397826954992
J1302−6313	2011-07-29-01:17:55/11	967.7817	532.26	10.6	24839.418691522	0.49663379788399	55770.802676492029
J1309−6526	2012-04-14-10:01:12/03	397.7251	360.68	8.0	27094.3648116048	5.97537183761597	56031.125057299541
J1327−6400	2011-12-07-17:28:25/11	280.7610	662.14	8.5	26336.5920056663	1.14667165279388	55902.588534666080
J1337−6306	2011-12-10-16:54:46/06	208.1117	770.54	5.0	31230.4609661705	3.99312567710876	55905.551022289394
J1429−5935	2012-02-15-22:43:23/12	763.8441	452.51	9.4	23484.8050807693	0.13974750041962	55972.683846828651
J1526−5633	2010-12-09-18:50:31/06	301.8818	328.42	8.9	25483.374940948	0	55539.534459834831
J1552−4937	2012-01-19-17:59:08/05	6.2839	113.85	11.3	25483.374940948	0	55945.498776964458
J1607−5140	2012-04-05-17:14:07/04	342.5404	561.30	6.0	25135.4672099139	2.04758715629578	56022.447319454142
J1610−5303	2012-08-02-09:33:26/03	786.3174	384.39	31.5	23544.2683172247	0.94431865215302	56141.142513330044
J1621−5039	2010-12-29-03:31:28/01	1084.0176	266.54	11.1	25476.8040597409	0.27327674627304	55559.058955496243
J1622−4950	2010-12-29-03:31:28/08	4285.9649	794.20	29.5	29872.686140058	5.93653869628906	55558.802423548157
J1624−4721	2011-05-09-12:46:41/12	448.7061	364.63	12.5	25483.374940948	0	55690.281797797790
J1626−4807	2010-12-31-21:03:27/11	293.8484	795.57	10.3	24606.5371443093	0.92786014080048	55561.604914473661
J1627−4706	2011-12-22-00:16:31/10	140.7716	464.19	8.7	33263.0986177139	1.06929671764374	55916.836951136618
J1632−4621	2011-06-26-07:39:22/09	1709.4977	535.49	12.8	35331.5627787543	1.14482700824738	55738.165229247381
J1633−4805	2011-12-30-19:34:06/13	711.4500	1145.65	10.9	41691.3852481281	5.70246410369873	55925.598278999729
J1635−4513	2012-04-05-18:27:22/09	1596.2830	419.37	25.9	31520.431292396	5.25761795043945	56022.605642167364
J1637−46	2012-04-05-18:27:22/04	493.1222	617.17	8.6	22979.0246812812	0.40881499648094	56022.647129696037
J1643−4522	2012-08-02-10:46:47/08	1348.2043	501.42	9.3	25016.8636087569	0.7531675696373	56141.340329424485
J1644−44	2011-07-03-10:59:47/11	173.9166	530.26	9.1	25483.374940948	0	55745.207561686686
J1644−46	2011-05-17-13:02:58/02	250.8684	404.78	6.0	23544.2683172247	0.94431865215302	55698.288022589302
J1644−4657	2011-05-17-13:02:58/08	125.9918	716.74	7.6	28112.2851796621	1.25053977966309	55698.408426516449
J1653−4854	2012-10-02-03:32:53/02	3059.0070	350.32	14.9	24606.5371443093	0.92786014080048	56201.875354288473
J1654−4245	2012-09-24-05:30:04/01	1101.9559	934.88	7.5	28112.2851796621	1.25053977966309	56194.093912627563

psrname	pointing/beam	period_topo	dm_topo	sigma	porb	asini	t0
J1709−4342	2011-12-12-05:16:53/02	1733.4461	291.94	13.1	39186.7677999458	8.69191551208496	55906.784259856293
J1737−3320	2013-01-30-18:19:00/11	816.6076	817.20	9.4	28455.8128049551	2.35764789581299	56322.626937362162
J1738−3107	2011-12-22-22:10:51/03	549.2362	747.44	10.1	24390.86317623	1.94023406505585	55917.670483210735
J1738−3316	2010-12-11-03:35:21/04	730.6016	263.73	7.7	23363.2652803823	1.01791310310364	55541.029786962987
J1755−25211	2013-01-22-01:43:12/08	1004.4311	821.92	9.0	23484.8050807693	0.13974750041962	56313.808719513836
J1755−2534	2012-07-26-12:48:29/02	233.5580	590.65	9.0	25483.374940948	0	56134.283047797791
J1801−2115	2012-03-30-22:36:23/01	437.9443	780.18	7.2	29546.1212438044	1.54342865943909	56016.937568404333
J1801−2154	2012-08-05-07:04:25/08	375.4066	392.04	10.0	33263.0986177139	1.06929671764374	56144.120215025512
J1805−2447	2011-05-07-14:31:02/10	661.3271	264.05	32.1	23484.8050807693	0.13974750041962	55688.341937106430
J1806−1618	2012-07-23-10:14:20/01	669.4311	313.64	6.1	30230.0360115379	8.2655611038208	56131.249968836913
J1806−1618	2012-07-23-10:14:20/01	668.3422	283.59	6.0	25483.374940948	0	56131.175999186686
J1808−1517	2013-04-02-18:57:44/10	544.6956	205.67	6.2	31043.8835104638	1.87641561031342	56384.587013440774
J1809−2004	2013-04-02-16:32:16/05	434.9209	922.21	7.3	28112.2851796621	1.25053977966309	56384.553773738669
J1818−1448	2013-04-02-20:10:28/09	281.5893	674.73	6.5	41691.3852481281	5.70246410369873	56384.623533629354
J1819−1131	2012-08-05-08:17:44/02	1387.4830	595.19	8.9	26232.8340110124	2.29141592979431	56144.067179769459
J1819−1717	2012-07-26-14:01:59/03	393.6209	409.68	11.0	33263.0986177139	1.06929671764374	56134.410191877359
J1822−1617	2012-07-25-14:25:48/12	830.2264	614.16	7.9	25901.6059270361	4.70340585708618	56133.325789535767
J1830−1313	2012-02-17-01:36:23/02	747.1243	543.76	7.7	25483.374940948	0	55973.816311686685
J1837−0604	2012-08-05-09:30:53/11	96.3159	455.95	22.7	25483.374940948	0	56144.145825575572
J1837−0822	2013-01-01-02:30:42/11	1097.8933	507.18	13.7	32379.2163759524	5.86167240142822	56292.749294886125
J1839−0321	2011-05-09-16:25:20/07	238.7646	454.27	24.4	25483.374940948	0	55690.433638075570
J1839−0627	2012-08-05-09:30:53/04	485.0384	101.29	8.0	33263.0986177139	1.06929671764374	56144.221927988474
J1839−1238	2011-01-26-00:34:08/04	1911.4144	212.46	11.5	25476.8040597409	0.27327674627304	55586.935807348091
J1840−0559	2012-08-05-09:30:53/05	859.3106	331.78	15.9	24839.418691522	0.49663379788399	56144.145014454989
J1840−0626	2012-08-05-09:30:53/10	1903.0043	690.56	27.6	26830.0569509955	6.73248815536499	56144.111760377891
J1840−0643	2012-08-05-09:30:53/03	35.5799	491.22	17.8	25483.374940948	0	56144.145825575572

APPENDIX B

Flux Density Limits for Normal Pulsars in a compact orbit around a Black Hole.

B.1 $P_{\text{spin}} = 100$ ms

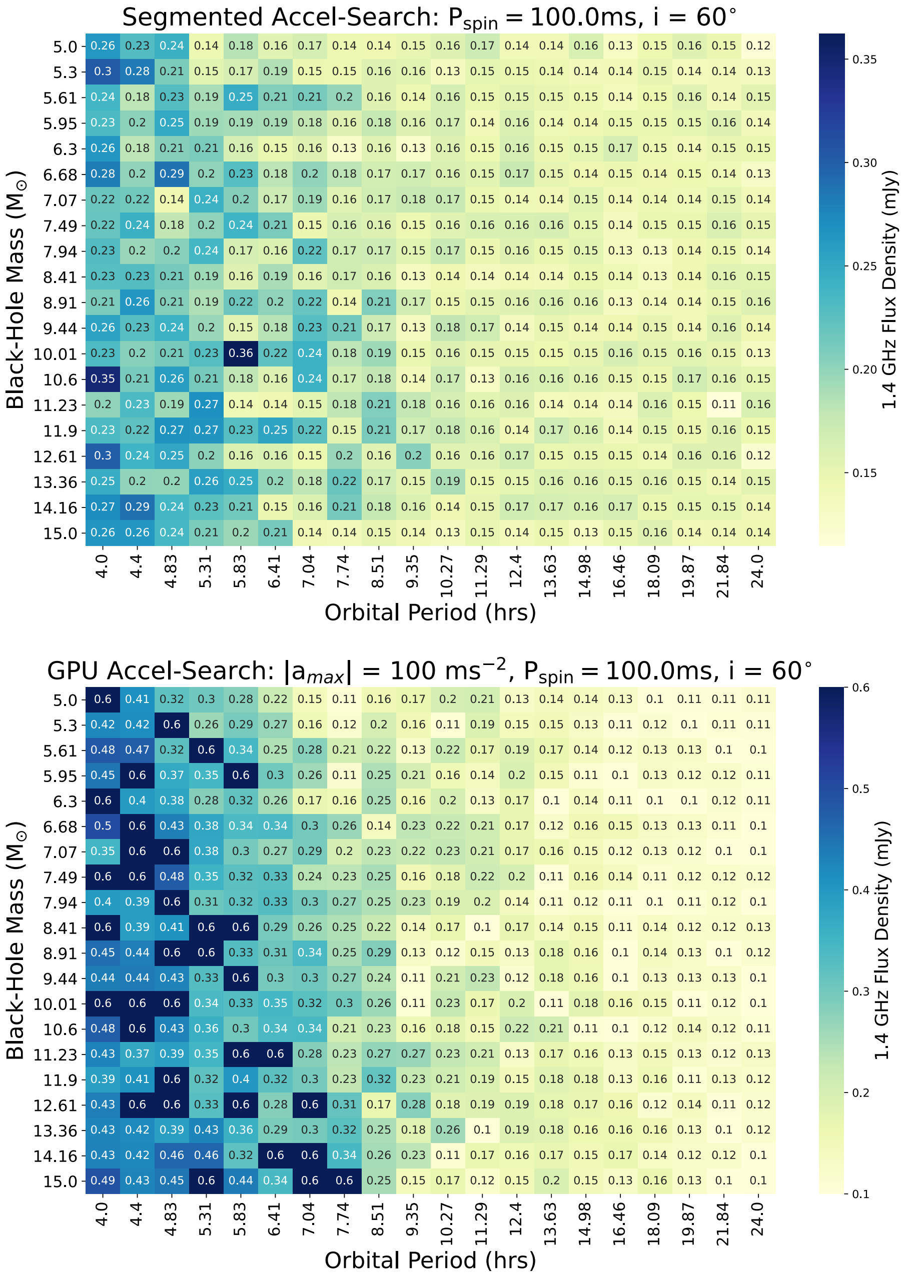
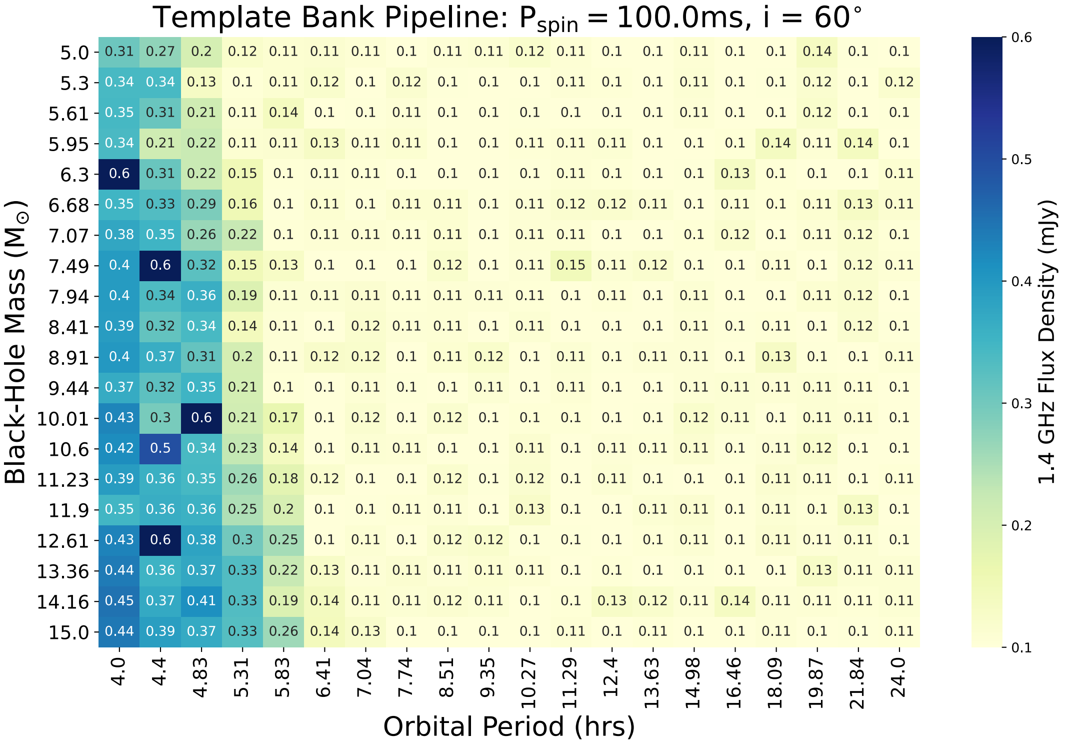
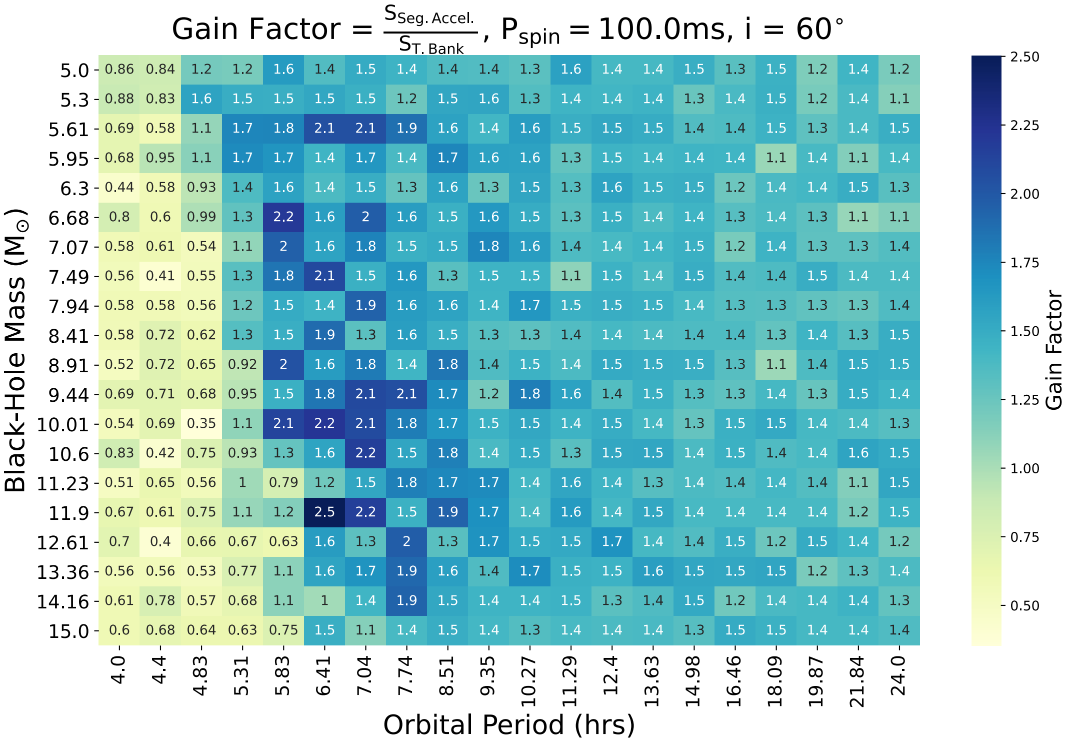


Figure B.1: Minimum detectable flux density limit for a 100 ms pulsar orbiting a black hole of varying mass as a function of the orbital period of the binary. In the top we show the results from a segmented acceleration search pipeline (Ng et al., 2015) and in the bottom we show the results from the ongoing reprocessing of LOWLAT using a GPU acceleration search (Sengar et al., prep). See See section 3.3.8 for full discussion.



	4.0	4.4	4.83	5.31	5.83	6.41	7.04	7.74	8.51	9.35	10.27	11.29	12.4	13.63	14.98	16.46	18.09	19.87	21.84	24.0
5.0	0.31	0.27	0.2	0.12	0.11	0.11	0.11	0.1	0.11	0.11	0.12	0.11	0.1	0.1	0.11	0.1	0.1	0.14	0.1	0.1
5.3	0.34	0.34	0.13	0.1	0.11	0.12	0.1	0.12	0.1	0.1	0.1	0.11	0.1	0.1	0.11	0.1	0.1	0.12	0.1	0.12
5.61	0.35	0.31	0.21	0.11	0.14	0.1	0.1	0.11	0.1	0.1	0.1	0.1	0.1	0.1	0.11	0.1	0.1	0.12	0.1	0.1
5.95	0.34	0.21	0.22	0.11	0.11	0.13	0.11	0.11	0.1	0.1	0.11	0.11	0.11	0.1	0.1	0.1	0.14	0.11	0.14	0.1
6.3	0.6	0.31	0.22	0.15	0.1	0.11	0.11	0.1	0.1	0.1	0.1	0.11	0.1	0.1	0.1	0.13	0.1	0.1	0.1	0.11
6.68	0.35	0.33	0.29	0.16	0.1	0.11	0.1	0.11	0.11	0.1	0.11	0.12	0.12	0.11	0.1	0.11	0.1	0.11	0.13	0.11
7.07	0.38	0.35	0.26	0.22	0.1	0.11	0.11	0.11	0.11	0.1	0.11	0.11	0.1	0.1	0.1	0.12	0.1	0.11	0.12	0.1
7.49	0.4	0.6	0.32	0.15	0.13	0.1	0.1	0.1	0.12	0.1	0.11	0.15	0.11	0.12	0.1	0.1	0.11	0.1	0.12	0.11
7.94	0.4	0.34	0.36	0.19	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.1	0.11	0.1	0.11	0.1	0.1	0.11	0.12	0.1
8.41	0.39	0.32	0.34	0.14	0.11	0.1	0.12	0.11	0.11	0.1	0.11	0.1	0.1	0.1	0.11	0.1	0.11	0.1	0.12	0.1
8.91	0.4	0.37	0.31	0.2	0.11	0.12	0.12	0.1	0.11	0.12	0.1	0.11	0.1	0.11	0.11	0.1	0.13	0.1	0.1	0.11
9.44	0.37	0.32	0.35	0.21	0.1	0.1	0.11	0.1	0.1	0.11	0.1	0.11	0.1	0.1	0.11	0.11	0.11	0.11	0.11	0.1
10.01	0.43	0.3	0.6	0.21	0.17	0.1	0.12	0.1	0.12	0.1	0.1	0.1	0.1	0.1	0.12	0.11	0.1	0.11	0.11	0.1
10.6	0.42	0.5	0.34	0.23	0.14	0.1	0.11	0.11	0.1	0.1	0.11	0.1	0.11	0.11	0.11	0.1	0.11	0.12	0.1	0.1
11.23	0.39	0.36	0.35	0.26	0.18	0.12	0.1	0.1	0.12	0.1	0.12	0.1	0.11	0.1	0.1	0.1	0.11	0.11	0.1	0.11
11.9	0.35	0.36	0.36	0.25	0.2	0.1	0.1	0.11	0.11	0.1	0.13	0.1	0.1	0.11	0.11	0.1	0.11	0.1	0.13	0.1
12.61	0.43	0.6	0.38	0.3	0.25	0.1	0.11	0.1	0.12	0.12	0.1	0.1	0.1	0.1	0.11	0.1	0.11	0.11	0.11	0.1
13.36	0.44	0.36	0.37	0.33	0.22	0.13	0.11	0.11	0.11	0.11	0.11	0.1	0.1	0.1	0.1	0.1	0.1	0.13	0.11	0.11
14.16	0.45	0.37	0.41	0.33	0.19	0.14	0.11	0.11	0.12	0.11	0.1	0.1	0.13	0.12	0.11	0.14	0.11	0.11	0.11	0.11
15.0	0.44	0.39	0.37	0.33	0.26	0.14	0.13	0.1	0.1	0.1	0.1	0.11	0.1	0.11	0.1	0.1	0.11	0.1	0.1	0.11



	4.0	4.4	4.83	5.31	5.83	6.41	7.04	7.74	8.51	9.35	10.27	11.29	12.4	13.63	14.98	16.46	18.09	19.87	21.84	24.0
5.0	0.86	0.84	1.2	1.2	1.6	1.4	1.5	1.4	1.4	1.4	1.3	1.6	1.4	1.4	1.5	1.3	1.5	1.2	1.4	1.2
5.3	0.88	0.83	1.6	1.5	1.5	1.5	1.5	1.2	1.5	1.6	1.3	1.4	1.4	1.4	1.3	1.4	1.5	1.2	1.4	1.1
5.61	0.69	0.58	1.1	1.7	1.8	2.1	2.1	1.9	1.6	1.4	1.6	1.5	1.5	1.5	1.4	1.4	1.5	1.3	1.4	1.5
5.95	0.68	0.95	1.1	1.7	1.7	1.4	1.7	1.4	1.7	1.6	1.6	1.3	1.5	1.4	1.4	1.4	1.1	1.4	1.1	1.4
6.3	0.44	0.58	0.93	1.4	1.6	1.4	1.5	1.3	1.6	1.3	1.5	1.3	1.6	1.5	1.5	1.2	1.4	1.4	1.5	1.3
6.68	0.8	0.6	0.99	1.3	2.2	1.6	2	1.6	1.5	1.6	1.5	1.3	1.5	1.4	1.4	1.3	1.4	1.3	1.1	1.1
7.07	0.58	0.61	0.54	1.1	2	1.6	1.8	1.5	1.5	1.8	1.6	1.4	1.4	1.4	1.5	1.2	1.4	1.3	1.3	1.4
7.49	0.56	0.41	0.55	1.3	1.8	2.1	1.5	1.6	1.3	1.5	1.5	1.1	1.5	1.4	1.5	1.4	1.4	1.5	1.4	1.4
7.94	0.58	0.58	0.56	1.2	1.5	1.4	1.9	1.6	1.6	1.4	1.7	1.5	1.5	1.5	1.4	1.3	1.3	1.3	1.3	1.4
8.41	0.58	0.72	0.62	1.3	1.5	1.9	1.3	1.6	1.5	1.3	1.3	1.4	1.4	1.4	1.4	1.4	1.3	1.4	1.3	1.5
8.91	0.52	0.72	0.65	0.92	2	1.6	1.8	1.4	1.8	1.4	1.5	1.4	1.5	1.5	1.5	1.3	1.1	1.4	1.5	1.5
9.44	0.69	0.71	0.68	0.95	1.5	1.8	2.1	2.1	1.7	1.2	1.8	1.6	1.4	1.5	1.3	1.3	1.4	1.3	1.5	1.4
10.01	0.54	0.69	0.35	1.1	2.1	2.2	2.1	1.8	1.7	1.5	1.5	1.4	1.5	1.4	1.3	1.5	1.5	1.4	1.4	1.3
10.6	0.83	0.42	0.75	0.93	1.3	1.6	2.2	1.5	1.8	1.4	1.5	1.3	1.5	1.5	1.4	1.5	1.4	1.4	1.6	1.5
11.23	0.51	0.65	0.56	1	0.79	1.2	1.5	1.8	1.7	1.7	1.4	1.6	1.4	1.3	1.4	1.4	1.4	1.4	1.1	1.5
11.9	0.67	0.61	0.75	1.1	1.2	2.5	2.2	1.5	1.9	1.7	1.4	1.6	1.4	1.5	1.4	1.4	1.4	1.4	1.2	1.5
12.61	0.7	0.4	0.66	0.67	0.63	1.6	1.3	2	1.3	1.7	1.5	1.5	1.7	1.4	1.4	1.5	1.2	1.5	1.4	1.2
13.36	0.56	0.56	0.53	0.77	1.1	1.6	1.7	1.9	1.6	1.4	1.7	1.5	1.5	1.6	1.5	1.5	1.5	1.2	1.3	1.4
14.16	0.61	0.78	0.57	0.68	1.1	1	1.4	1.9	1.5	1.4	1.4	1.5	1.3	1.4	1.5	1.2	1.4	1.4	1.4	1.3
15.0	0.6	0.68	0.64	0.63	0.75	1.5	1.1	1.4	1.5	1.4	1.3	1.4	1.4	1.4	1.3	1.5	1.5	1.4	1.4	1.4

Figure B.2: Minimum detectable flux density limit for a 100 ms pulsar orbiting a black hole of varying mass as a function of the orbital period of the binary. In the top we show the results from the template-bank pipeline which is presented in this work and in the bottom we show the sensitivity gain factor compared to the best-case acceleration-search pipeline i.e the segmented acceleration search pipeline (Ng et al., 2015). A factor above 1 indicates improvement from the template-bank pipeline and conversely below 1 indicates better performance from the acceleration-search pipeline. See See section 3.3.8 for full discussion.

B.2 $P_{\rm spin}$ = 1000 ms

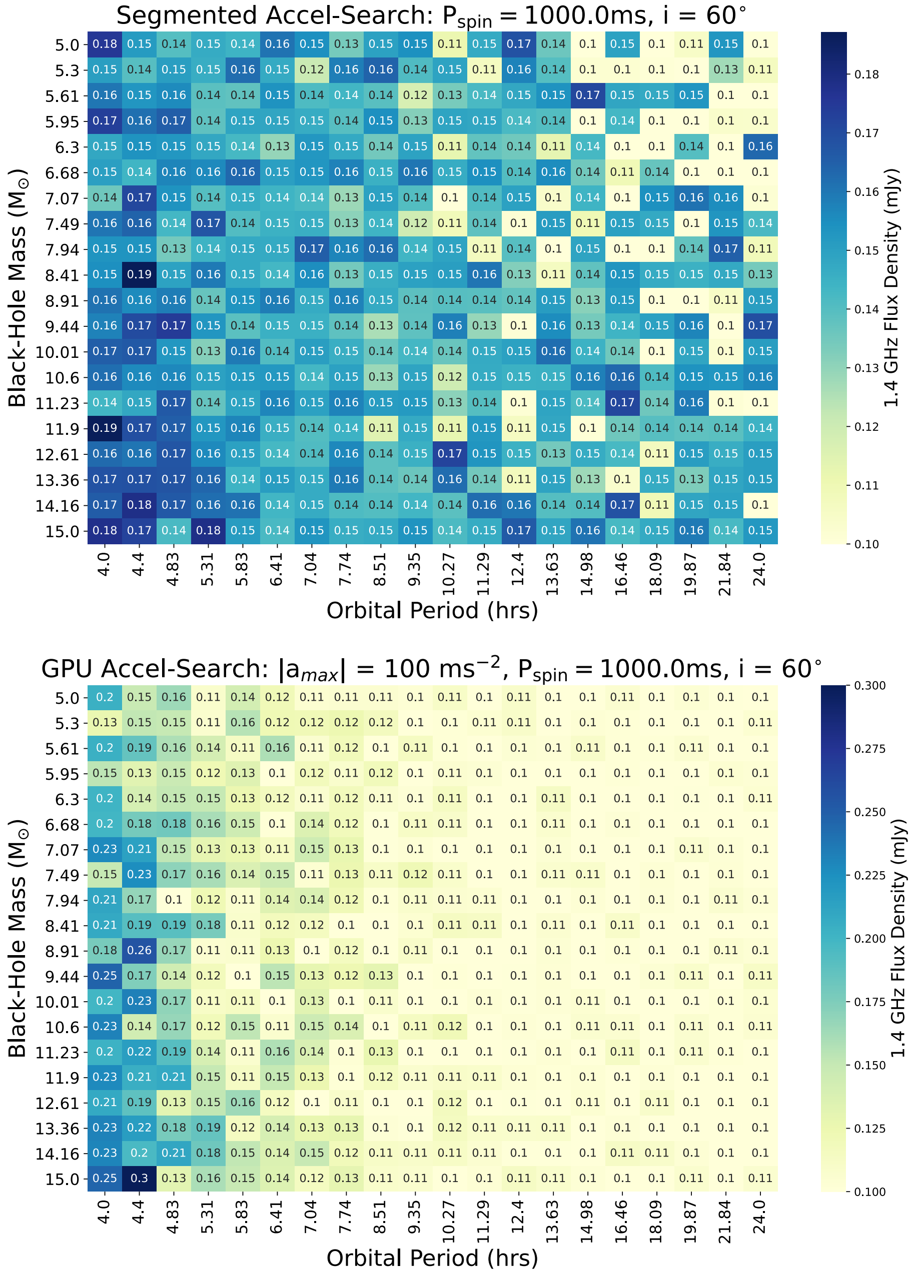


Figure B.3: Minimum detectable flux density limit for a 1000 ms pulsar orbiting a black hole of varying mass as a function of the orbital period of the binary. In the top we show the results from a segmented acceleration search pipeline (Ng et al., 2015) and in the bottom we show the results from the ongoing reprocessing of LOWLAT using a GPU acceleration search (Sengar et al., prep). See section 3.3.8 for full discussion.

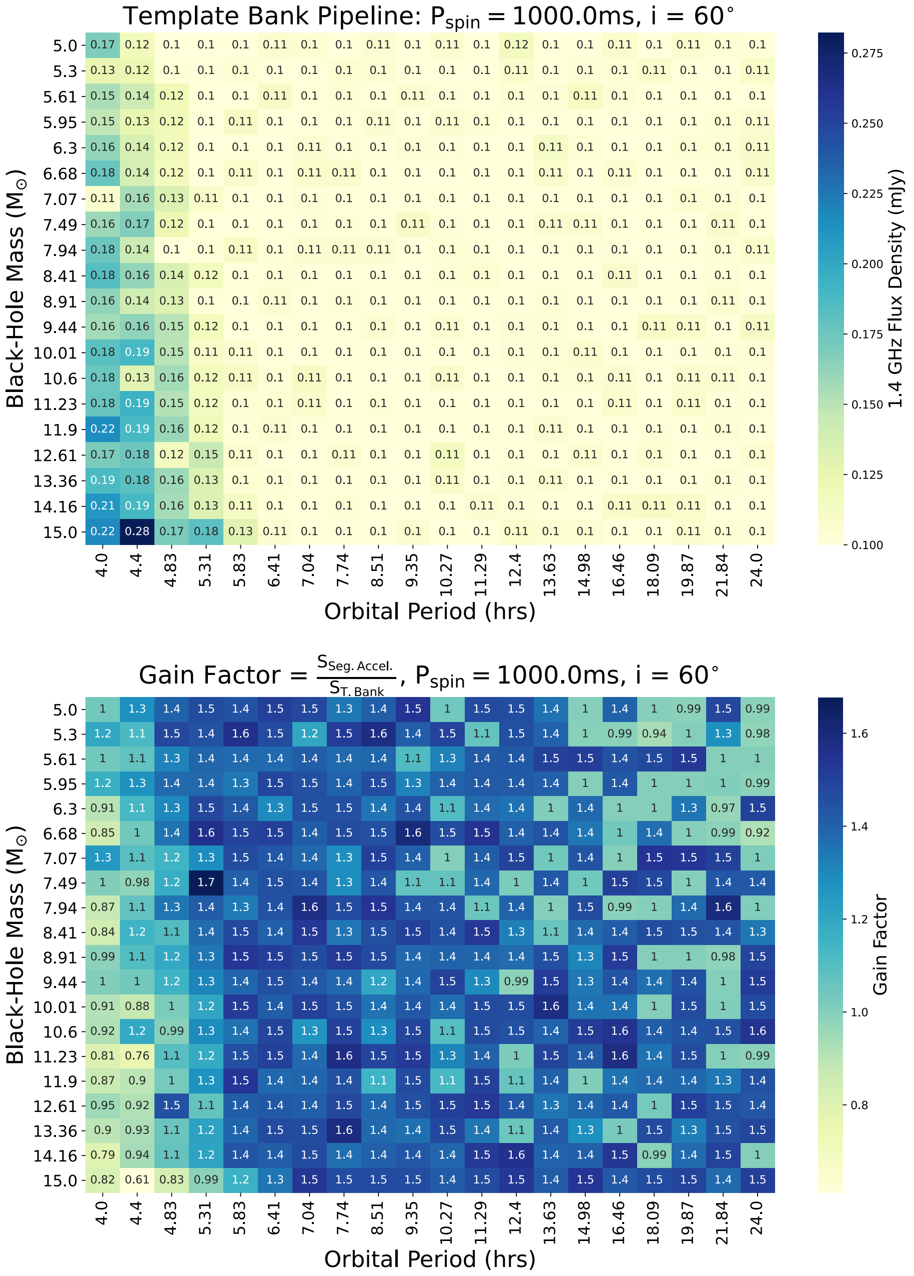
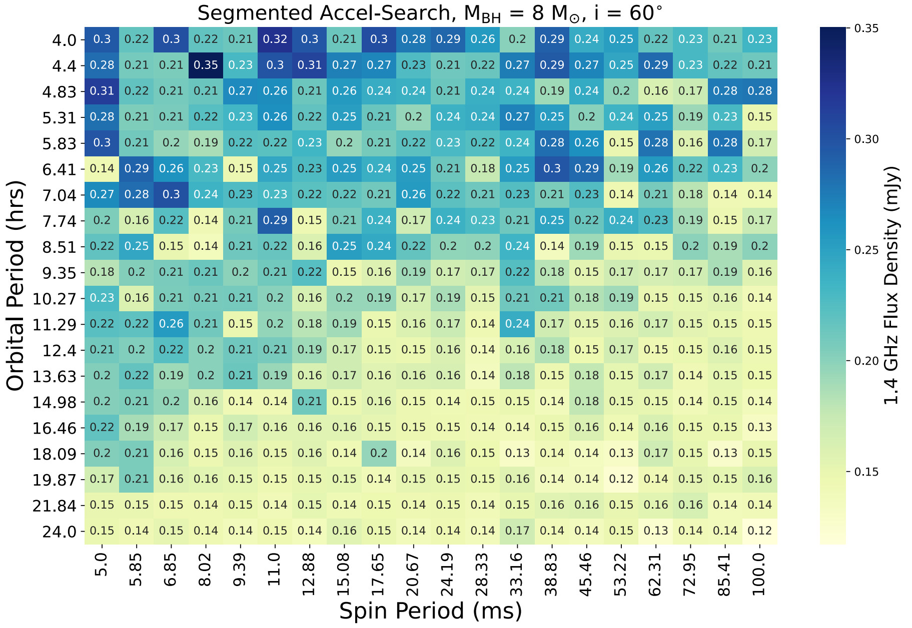


Figure B.4: Minimum detectable flux density limit for a 1000 ms pulsar orbiting a black hole of varying mass as a function of the orbital period of the binary. In the top we show the results from the template-bank pipeline which is presented in this work and in the bottom we show the sensitivity gain factor compared to the best-case acceleration-search pipeline i.e the segmented acceleration search pipeline (Ng et al., 2015). A factor above 1 indicates improvement from the template-bank pipeline and conversely below 1 indicates better performance from the acceleration-search pipeline. See section 3.3.8 for full discussion.

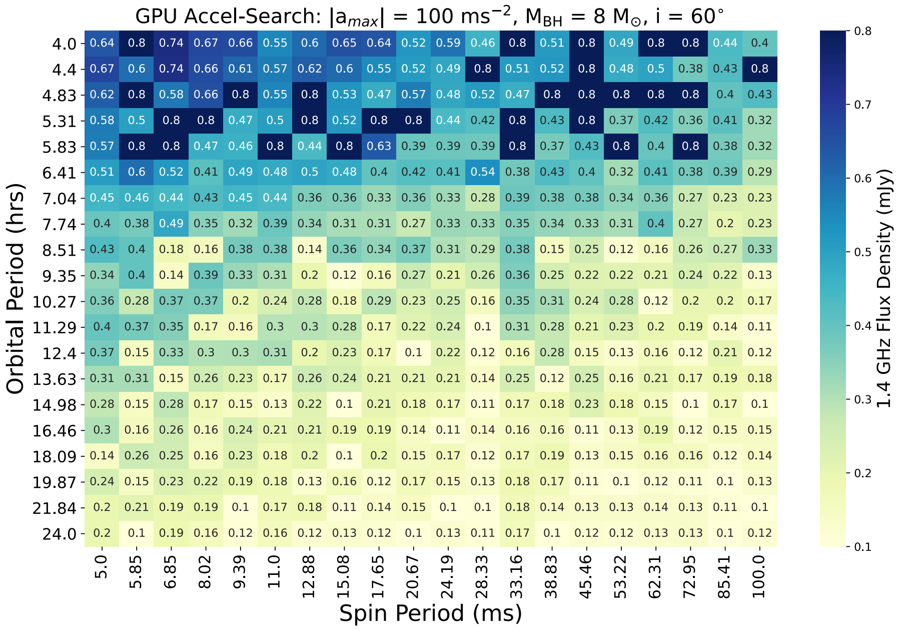
Appendix C

Flux Density Limits for PSR-BH binaries from past searches in LOWLAT in the P_{orb}-P_{spin} plane.

C.1 Sensitivity of acceleration-search pipelines.



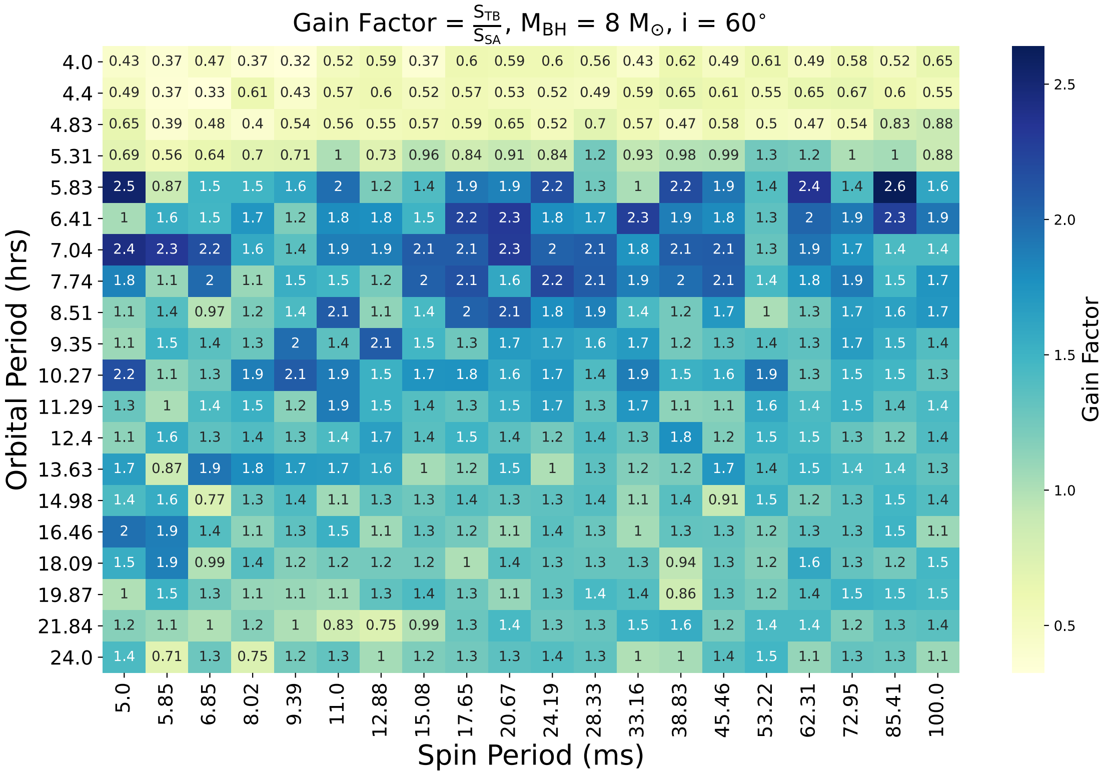
	5.0	5.85	6.85	8.02	9.39	11.0	12.88	15.08	17.65	20.67	24.19	28.33	33.16	38.83	45.46	53.22	62.31	72.95	85.41	100.0
4.0	0.3	0.22	0.3	0.22	0.21	0.32	0.3	0.21	0.3	0.28	0.29	0.26	0.2	0.29	0.24	0.25	0.22	0.23	0.21	0.23
4.4	0.28	0.21	0.21	0.35	0.23	0.3	0.31	0.27	0.27	0.23	0.21	0.22	0.27	0.29	0.27	0.25	0.29	0.23	0.22	0.21
4.83	0.31	0.22	0.21	0.21	0.27	0.26	0.21	0.26	0.24	0.24	0.21	0.24	0.24	0.19	0.24	0.2	0.16	0.17	0.28	0.28
5.31	0.28	0.21	0.21	0.22	0.23	0.26	0.22	0.25	0.21	0.2	0.24	0.24	0.27	0.25	0.2	0.24	0.25	0.19	0.23	0.15
5.83	0.3	0.21	0.2	0.19	0.22	0.22	0.23	0.2	0.21	0.22	0.23	0.22	0.24	0.28	0.26	0.15	0.28	0.16	0.28	0.17
6.41	0.14	0.29	0.26	0.23	0.15	0.25	0.23	0.25	0.24	0.25	0.21	0.18	0.25	0.3	0.29	0.19	0.26	0.22	0.23	0.2
7.04	0.27	0.28	0.3	0.24	0.23	0.23	0.22	0.22	0.21	0.26	0.22	0.21	0.23	0.21	0.23	0.14	0.21	0.18	0.14	0.14
7.74	0.2	0.16	0.22	0.14	0.21	0.29	0.15	0.21	0.24	0.17	0.24	0.23	0.21	0.25	0.22	0.24	0.23	0.19	0.15	0.17
8.51	0.22	0.25	0.15	0.14	0.21	0.22	0.16	0.25	0.24	0.22	0.2	0.2	0.24	0.14	0.19	0.15	0.15	0.2	0.19	0.2
9.35	0.18	0.2	0.21	0.21	0.2	0.21	0.22	0.15	0.16	0.19	0.17	0.17	0.22	0.18	0.15	0.17	0.17	0.17	0.19	0.16
10.27	0.23	0.16	0.21	0.21	0.21	0.2	0.16	0.2	0.19	0.17	0.19	0.15	0.21	0.21	0.18	0.19	0.15	0.15	0.16	0.14
11.29	0.22	0.22	0.26	0.21	0.15	0.2	0.18	0.19	0.15	0.16	0.17	0.14	0.24	0.17	0.15	0.16	0.17	0.15	0.15	0.15
12.4	0.21	0.2	0.22	0.2	0.21	0.21	0.19	0.17	0.15	0.15	0.16	0.14	0.16	0.18	0.15	0.17	0.15	0.15	0.16	0.15
13.63	0.2	0.22	0.19	0.2	0.21	0.19	0.16	0.17	0.16	0.16	0.16	0.14	0.18	0.15	0.18	0.15	0.17	0.14	0.15	0.15
14.98	0.2	0.21	0.2	0.16	0.14	0.14	0.21	0.15	0.16	0.15	0.14	0.15	0.15	0.14	0.18	0.15	0.15	0.14	0.15	0.14
16.46	0.22	0.19	0.17	0.15	0.17	0.16	0.16	0.16	0.15	0.15	0.15	0.14	0.14	0.15	0.16	0.14	0.16	0.15	0.15	0.13
18.09	0.2	0.21	0.16	0.15	0.16	0.15	0.16	0.14	0.2	0.14	0.16	0.15	0.13	0.14	0.14	0.13	0.17	0.15	0.13	0.15
19.87	0.17	0.21	0.16	0.16	0.15	0.15	0.15	0.15	0.14	0.15	0.15	0.15	0.16	0.14	0.14	0.12	0.14	0.15	0.15	0.16
21.84	0.15	0.15	0.15	0.14	0.15	0.14	0.15	0.14	0.15	0.14	0.15	0.14	0.15	0.16	0.16	0.15	0.16	0.16	0.14	0.14
24.0	0.15	0.14	0.15	0.14	0.14	0.15	0.14	0.16	0.15	0.14	0.14	0.14	0.17	0.14	0.14	0.15	0.13	0.14	0.14	0.12



	5.0	5.85	6.85	8.02	9.39	11.0	12.88	15.08	17.65	20.67	24.19	28.33	33.16	38.83	45.46	53.22	62.31	72.95	85.41	100.0
4.0	0.64	0.8	0.74	0.67	0.66	0.55	0.6	0.65	0.64	0.52	0.59	0.46	0.8	0.51	0.8	0.49	0.8	0.8	0.44	0.4
4.4	0.67	0.6	0.74	0.66	0.61	0.57	0.62	0.6	0.55	0.52	0.49	0.8	0.51	0.52	0.8	0.48	0.5	0.38	0.43	0.8
4.83	0.62	0.8	0.58	0.66	0.8	0.55	0.8	0.53	0.47	0.57	0.48	0.52	0.47	0.8	0.8	0.8	0.8	0.8	0.4	0.43
5.31	0.58	0.5	0.8	0.8	0.47	0.5	0.8	0.52	0.8	0.8	0.44	0.42	0.8	0.43	0.8	0.37	0.42	0.36	0.41	0.32
5.83	0.57	0.8	0.8	0.47	0.46	0.8	0.44	0.8	0.63	0.39	0.39	0.39	0.8	0.37	0.43	0.8	0.4	0.8	0.38	0.32
6.41	0.51	0.6	0.52	0.41	0.49	0.48	0.5	0.48	0.4	0.42	0.41	0.54	0.38	0.43	0.4	0.32	0.41	0.38	0.39	0.29
7.04	0.45	0.46	0.44	0.43	0.45	0.44	0.36	0.36	0.33	0.36	0.33	0.28	0.39	0.38	0.38	0.34	0.36	0.27	0.23	0.23
7.74	0.4	0.38	0.49	0.35	0.32	0.39	0.34	0.31	0.31	0.27	0.33	0.33	0.35	0.34	0.33	0.31	0.4	0.27	0.2	0.23
8.51	0.43	0.4	0.18	0.16	0.38	0.38	0.14	0.36	0.34	0.37	0.31	0.29	0.38	0.15	0.25	0.12	0.16	0.26	0.27	0.33
9.35	0.34	0.4	0.14	0.39	0.33	0.31	0.2	0.12	0.16	0.27	0.21	0.26	0.36	0.25	0.22	0.22	0.21	0.24	0.22	0.13
10.27	0.36	0.28	0.37	0.37	0.2	0.24	0.28	0.18	0.29	0.23	0.25	0.16	0.35	0.31	0.24	0.28	0.12	0.2	0.2	0.17
11.29	0.4	0.37	0.35	0.17	0.16	0.3	0.3	0.28	0.17	0.22	0.24	0.1	0.31	0.28	0.21	0.23	0.2	0.19	0.14	0.11
12.4	0.37	0.15	0.33	0.3	0.3	0.31	0.2	0.23	0.17	0.1	0.22	0.12	0.16	0.28	0.15	0.13	0.16	0.12	0.21	0.12
13.63	0.31	0.31	0.15	0.26	0.23	0.17	0.26	0.24	0.21	0.21	0.21	0.14	0.25	0.12	0.25	0.16	0.21	0.17	0.19	0.18
14.98	0.28	0.15	0.28	0.17	0.15	0.13	0.22	0.1	0.21	0.18	0.17	0.11	0.17	0.18	0.23	0.18	0.15	0.1	0.17	0.1
16.46	0.3	0.16	0.26	0.16	0.24	0.21	0.21	0.19	0.19	0.14	0.11	0.14	0.16	0.16	0.11	0.13	0.19	0.12	0.15	0.15
18.09	0.14	0.26	0.25	0.16	0.23	0.18	0.2	0.1	0.2	0.15	0.17	0.12	0.17	0.19	0.13	0.15	0.16	0.16	0.12	0.14
19.87	0.24	0.15	0.23	0.22	0.19	0.18	0.13	0.16	0.12	0.17	0.15	0.13	0.18	0.17	0.11	0.1	0.12	0.11	0.1	0.13
21.84	0.2	0.21	0.19	0.19	0.1	0.17	0.18	0.11	0.14	0.15	0.1	0.1	0.18	0.14	0.13	0.13	0.14	0.13	0.11	0.1
24.0	0.2	0.1	0.19	0.16	0.12	0.16	0.12	0.13	0.12	0.1	0.13	0.11	0.17	0.1	0.12	0.12	0.13	0.13	0.1	0.12

Figure C.1: Minimum detectable flux density limit for a PSR orbiting a black hole of mass 8 $M_\odot$ as a function of orbital period and spin-period of the pulsar. In the top we show the results from the segmented acceleration search pipeline (Ng et al., 2015) and in the bottom we show the results from the ongoing GPU accel. search reprocessing of LOWLAT in the full-length 72-minute observations (Sengar et al., prep). See section 3.3.9 for full discussion.

C.2 Gain Factor from the template-bank pipeline in the P_{orb}-P_{spin} plane.



	5.0	5.85	6.85	8.02	9.39	11.0	12.88	15.08	17.65	20.67	24.19	28.33	33.16	38.83	45.46	53.22	62.31	72.95	85.41	100.0
4.0	0.43	0.37	0.47	0.37	0.32	0.52	0.59	0.37	0.6	0.59	0.6	0.56	0.43	0.62	0.49	0.61	0.49	0.58	0.52	0.65
4.4	0.49	0.37	0.33	0.61	0.43	0.57	0.6	0.52	0.57	0.53	0.52	0.49	0.59	0.65	0.61	0.55	0.65	0.67	0.6	0.55
4.83	0.65	0.39	0.48	0.4	0.54	0.56	0.55	0.57	0.59	0.65	0.52	0.7	0.57	0.47	0.58	0.5	0.47	0.54	0.83	0.88
5.31	0.69	0.56	0.64	0.7	0.71	1	0.73	0.96	0.84	0.91	0.84	1.2	0.93	0.98	0.99	1.3	1.2	1	1	0.88
5.83	2.5	0.87	1.5	1.5	1.6	2	1.2	1.4	1.9	1.9	2.2	1.3	1	2.2	1.9	1.4	2.4	1.4	2.6	1.6
6.41	1	1.6	1.5	1.7	1.2	1.8	1.8	1.5	2.2	2.3	1.8	1.7	2.3	1.9	1.8	1.3	2	1.9	2.3	1.9
7.04	2.4	2.3	2.2	1.6	1.4	1.9	1.9	2.1	2.1	2.3	2	2.1	1.8	2.1	2.1	1.3	1.9	1.7	1.4	1.4
7.74	1.8	1.1	2	1.1	1.5	1.5	1.2	2	2.1	1.6	2.2	2.1	1.9	2	2.1	1.4	1.8	1.9	1.5	1.7
8.51	1.1	1.4	0.97	1.2	1.4	2.1	1.1	1.4	2	2.1	1.8	1.9	1.4	1.2	1.7	1	1.3	1.7	1.6	1.7
9.35	1.1	1.5	1.4	1.3	2	1.4	2.1	1.5	1.3	1.7	1.7	1.6	1.7	1.2	1.3	1.4	1.3	1.7	1.5	1.4
10.27	2.2	1.1	1.3	1.9	2.1	1.9	1.5	1.7	1.8	1.6	1.7	1.4	1.9	1.5	1.6	1.9	1.3	1.5	1.5	1.3
11.29	1.3	1	1.4	1.5	1.2	1.9	1.5	1.4	1.3	1.5	1.7	1.3	1.7	1.1	1.1	1.6	1.4	1.5	1.4	1.4
12.4	1.1	1.6	1.3	1.4	1.3	1.4	1.7	1.4	1.5	1.4	1.2	1.4	1.3	1.8	1.2	1.5	1.5	1.3	1.2	1.4
13.63	1.7	0.87	1.9	1.8	1.7	1.7	1.6	1	1.2	1.5	1	1.3	1.2	1.2	1.7	1.4	1.5	1.4	1.4	1.3
14.98	1.4	1.6	0.77	1.3	1.4	1.1	1.3	1.3	1.4	1.3	1.3	1.4	1.1	1.4	0.91	1.5	1.2	1.3	1.5	1.4
16.46	2	1.9	1.4	1.1	1.3	1.5	1.1	1.3	1.2	1.1	1.4	1.3	1	1.3	1.3	1.2	1.3	1.3	1.5	1.1
18.09	1.5	1.9	0.99	1.4	1.2	1.2	1.2	1.2	1	1.4	1.3	1.3	1.3	0.94	1.3	1.2	1.6	1.3	1.2	1.5
19.87	1	1.5	1.3	1.1	1.1	1.1	1.3	1.4	1.3	1.1	1.3	1.4	1.4	0.86	1.3	1.2	1.4	1.5	1.5	1.5
21.84	1.2	1.1	1	1.2	1	0.83	0.75	0.99	1.3	1.4	1.3	1.3	1.5	1.6	1.2	1.4	1.4	1.2	1.3	1.4
24.0	1.4	0.71	1.3	0.75	1.2	1.3	1	1.2	1.3	1.3	1.4	1.3	1	1	1.4	1.5	1.1	1.3	1.3	1.1

Figure C.2: Gain Factor from the template-bank pipeline in minimum detectable flux density towards PSR+BH binaries in the P_{orb}-P_{spin} plane assuming a fixed black mass of 8 $M_\odot$. A factor above 1 indicates improvement from the template-bank pipeline and conversely below 1 indicates better performance from the acceleration-search pipeline. See section 3.3.9 for full discussion.

Appendix D

Detailed Performance Comparison of SGAN, Re-trained PICS and Supervised Baseline Model.

Table D.1: Detailed performance comparison of Semi-Supervised GAN and the our Ensemble Supervised Machine Learning Baseline Model for varying amounts of labelled data. Each row represents the mean performance of the network after averaging across five different batches of labelled data.

			Accuracy		F-Score		FPR		G-Mean		Precision		Recall		Specificity	
			median	std	median	std	median	std	median	std	median	std	median	std	median	std
Labelled Samples	Unlabelled Samples	neural_net														
100	0	PICS Retrained	0.778114	0.021186	0.732734	0.035151	0.060684	0.014769	0.759390	0.028041	0.905080	0.015769	0.603222	0.050334	0.939316	0.014769
		Supervised Baseline	0.811978	0.030802	0.827461	0.027541	0.234638	0.057758	0.807474	0.031681	0.784844	0.043643	0.856594	0.030882	0.765362	0.057758
	5000	SGAN	0.947764	0.011199	0.948811	0.011012	0.064780	0.012405	0.947577	0.011209	0.937786	0.011890	0.959910	0.011033	0.935220	0.012405
500	0	PICS Retrained	0.870921	0.026020	0.858840	0.030660	0.035343	0.005588	0.866737	0.027780	0.957283	0.007754	0.783908	0.048144	0.964657	0.005588
		Supervised Baseline	0.925943	0.019918	0.927882	0.020159	0.093265	0.014035	0.925586	0.019763	0.911531	0.014496	0.944830	0.027332	0.906735	0.014035
	10000	SGAN	0.962169	0.003598	0.962963	0.003699	0.041726	0.006985	0.961964	0.003566	0.959504	0.006289	0.972087	0.009891	0.958274	0.006985
1000	0	PICS Retrained	0.915647	0.005872	0.910673	0.006351	0.027532	0.006123	0.913985	0.005989	0.969684	0.006642	0.859123	0.009233	0.972468	0.006123
		Supervised Baseline	0.945686	0.004898	0.947000	0.004645	0.070020	0.009455	0.945394	0.004990	0.932637	0.008242	0.963188	0.005629	0.929980	0.009455
	20000	SGAN	0.969537	0.002386	0.970052	0.002626	0.039535	0.007066	0.969419	0.002255	0.961790	0.006225	0.978456	0.011495	0.960465	0.007066
10000	0	PICS Retrained	0.970576	0.001187	0.970700	0.001225	0.025436	0.001286	0.970601	0.001182	0.974780	0.001203	0.966842	0.002887	0.974564	0.001286
		Supervised Baseline	0.976527	0.001230	0.976850	0.001208	0.032867	0.002169	0.976462	0.001232	0.968276	0.002017	0.986043	0.002044	0.967133	0.002169
	20000	SGAN	0.983989	0.000400	0.984271	0.000387	0.025817	0.001091	0.983861	0.000407	0.975074	0.001011	0.994099	0.000755	0.974183	0.001091
30000	0	PICS Retrained	0.973646	0.000919	0.973827	0.000905	0.025055	0.001558	0.973656	0.000921	0.975291	0.001504	0.971525	0.001054	0.974945	0.001558
		Supervised Baseline	0.982430	0.000326	0.982686	0.000305	0.024102	0.001742	0.982354	0.000340	0.976596	0.001623	0.988854	0.001309	0.975898	0.001742
	20000	SGAN	0.985973	0.001134	0.986206	0.001106	0.022673	0.001933	0.985863	0.001146	0.978075	0.001838	0.994005	0.000566	0.977327	0.001933
50814	0	PICS Retrained	0.973693	0.000928	0.973829	0.000904	0.023245	0.002171	0.973714	0.000933	0.976996	0.002072	0.970963	0.001352	0.976755	0.002171
		Supervised Baseline	0.983847	0.001124	0.984099	0.001093	0.023721	0.002387	0.983755	0.001138	0.977013	0.002252	0.991289	0.001282	0.976279	0.002387
	265172	SGAN	0.989468	0.004313	0.989614	0.004251	0.016290	0.004592	0.989403	0.004317	0.984159	0.004456	0.994661	0.004115	0.983710	0.004592

Table D.2: Performance comparison of SGAN, retrained version of PICS and our Supervised Baseline Model for different ranges of detection sigma values.

		FPR	Total Candidates	Precision	Sigma Avg	Sigma Max	Sigma Min	Recall	Specificity
Percentile	Model								
10	PICS Retrained	0.034501	2114	0.576923	3.700	7.40	0.00	0.447761	0.965499
	SGAN	0.021432	2114	0.823276	3.700	7.40	0.00	0.950249	0.978568
	Supervised Baseline	0.040774	2114	0.700000	3.700	7.40	0.00	0.905473	0.959226
20	PICS Retrained	0.081731	2107	0.943381	9.750	12.10	7.40	0.888627	0.918269
	SGAN	0.052885	2107	0.966206	9.750	12.10	7.40	0.986667	0.947115
	Supervised Baseline	0.055288	2107	0.963808	9.750	12.10	7.40	0.960784	0.944712
30	PICS Retrained	0.027174	2107	0.985465	14.550	17.00	12.10	0.989059	0.972826
	SGAN	0.014946	2107	0.992035	14.550	17.00	12.10	0.999271	0.985054
	Supervised Baseline	0.019022	2107	0.989766	14.550	17.00	12.10	0.987600	0.980978
40	PICS Retrained	0.016985	2117	0.986498	19.450	21.90	17.00	0.994894	0.983015
	SGAN	0.004246	2117	0.996596	19.450	21.90	17.00	0.996596	0.995754
	Supervised Baseline	0.010616	2117	0.991468	19.450	21.90	17.00	0.988936	0.989384
50	PICS Retrained	0.016765	2110	0.984657	24.500	27.10	21.90	0.995438	0.983235
	SGAN	0.007890	2110	0.992754	24.500	27.10	21.90	1.000000	0.992110
	Supervised Baseline	0.010848	2110	0.990009	24.500	27.10	21.90	0.994526	0.989152
60	PICS Retrained	0.009314	2116	0.988285	29.800	32.50	27.10	0.992513	0.990686
	SGAN	0.004234	2116	0.994681	29.800	32.50	27.10	1.000000	0.995766
	Supervised Baseline	0.009314	2116	0.988285	29.800	32.50	27.10	0.992513	0.990686
70	PICS Retrained	0.004095	2149	0.994624	36.185	39.87	32.50	0.996767	0.995905
	SGAN	0.001638	2149	0.997849	36.185	39.87	32.50	1.000000	0.998362
	Supervised Baseline	0.004914	2149	0.993562	36.185	39.87	32.50	0.997845	0.995086
80	PICS Retrained	0.013174	2111	0.991413	46.785	53.70	39.87	0.995298	0.986826
	SGAN	0.005988	2111	0.996097	46.785	53.70	39.87	1.000000	0.994012
	Supervised Baseline	0.015569	2111	0.989907	46.785	53.70	39.87	0.999216	0.984431
90	PICS Retrained	0.034483	2123	0.991860	74.795	95.89	53.70	0.993593	0.965517
	SGAN	0.029557	2123	0.993060	74.795	95.89	53.70	1.000000	0.970443
	Supervised Baseline	0.036946	2123	0.991324	74.795	95.89	53.70	0.998253	0.963054
100	PICS Retrained	0.007057	2118	0.985856	510.545	925.20	95.89	0.994294	0.992943
	SGAN	0.007763	2118	0.984551	510.545	925.20	95.89	1.000000	0.992237
	Supervised Baseline	0.009880	2118	0.980420	510.545	925.20	95.89	1.000000	0.990120

Appendix E

Coefficients of the taylor expansion of Eccentric Anomaly

Below, for completeness we mention the coefficients C_k, S_k taken from (Taff, 1985) and used in equations 5.15 and 5.16. These expressions were also given in Appendix A of (Dhurandhar & Vecchio, 2001). A derivation for these expressions can also be found in Appendix C of (Nieder et al., 2020).

$$\begin{aligned}
C_0 &= -\frac{1}{2}e, \\
C_1 &= 1 - \frac{3}{8}e^2 + \frac{5}{192}e^4 - \frac{7}{9216}e^6, \\
C_2 &= \frac{1}{2}e - \frac{1}{3}e^3 + \frac{1}{16}e^5, \\
C_3 &= \frac{3}{8}e^2 - \frac{45}{128}e^4 + \frac{567}{5120}e^6, \\
C_4 &= \frac{1}{3}e^3 - \frac{2}{5}e^5, \\
C_5 &= \frac{125}{384}e^4 - \frac{4375}{9216}e^6, \\
C_6 &= \frac{27}{80}e^5, \\
C_7 &= \frac{16807}{46080}e^6,
\end{aligned}$$

$$\begin{aligned}
S_1 &= 1 - \frac{5}{8}e^2 - \frac{11}{192}e^4 - \frac{457}{9216}e^6, \\
S_2 &= \frac{1}{2}e - \frac{5}{12}e^3 + \frac{1}{24}e^5, \\
S_3 &= \frac{3}{8}e^2 - \frac{51}{128}e^4 + \frac{543}{5120}e^6, \\
S_4 &= \frac{1}{3}e^3 - \frac{13}{30}e^5, \\
S_5 &= \frac{125}{384}e^4 - \frac{4625}{9216}e^6, \\
S_6 &= \frac{27}{80}e^5, \\
S_7 &= \frac{16807}{46080}e^6.
\end{aligned}$$

Bibliography

Abadi M., Agarwal A., Barham P., Brevdo E., Chen Z., Citro C., Corrado G. S., Davis A., Dean J., Devin M., Ghemawat S., Goodfellow I., Harp A., Irving G., Isard M., Jia Y., Jozefowicz R., Kaiser L., Kudlur M., Levenberg J., Mané D., Monga R., Moore S., Murray D., Olah C., Schuster M., Shlens J., Steiner B., Sutskever I., Talwar K., Tucker P., Vanhoucke V., Vasudevan V., Viégas F., Vinyals O., Warden P., Wattenberg M., Wicke M., Yu Y., Zheng X., , 2015, TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems (Cited on pages 136 and 147.)

Abbot B. P., et al., 2007, Phys. Rev. D, 76, 082001 (Cited on pages 56, 62 and 153.)

Abbot B. P., et al., 2009, Phys. Rev. D, 80, 042003 (Cited on pages 56, 62 and 153.)

Abbot B. P., et al., 2016, Phys. Rev. Lett., 116, 061102 (Cited on page 27.)

Abbot B. P., et al., 2017, Phys. Rev. Lett., 119, 161101 (Cited on page 27.)

Abbot B. P., et al., 2018, Phys. Rev. Lett., 121, 161101 (Cited on page 5.)

Ables J. G., Manchester R. N., 1976, A&A, 50, 177 (Cited on page 10.)

Agarwal D., Aggarwal K., Burke-Spolaor S., Lorimer D. R., Garver-Daniels N., 2019, arXiv e-prints, p. arXiv:1902.06343 (Cited on page 146.)

Allen B., Knispel B., Cordes J. M., Deneva J. S., Hessels J. W. T., Anderson D., Aulbert C., Bock O., Brazier A., Chatterjee S., et al. 2013, The Astrophysical Journal, 773, 91 (Cited on pages 42, 55, 62, 65, 151, 156, 158, 174, 177 and 181.)

Allen B., Papa M. A., Schutz B. F., 2002, Phys. Rev. D, 66, 102003 (Cited on pages 56 and 153.)

Alpar M. A., Cheng A. F., Ruderman M. A., Shaham J., 1982, Nature, 300, 728 (Cited on page 16.)

Andersen B. C., Ransom S. M., 2018, ApJ, 863, L13 (Cited on pages 42, 151 and 158.)

Anderson P. W., Itoh N., 1975, Nature, 256, 25 (Cited on page 27.)

Antoniadis J., Freire P. C. C., Wex N., Tauris T. M., Lynch R. S., van Kerkwijk M. H., Kramer M., Bassa C., Dhillon V. S., Driebe T., Hessels J. W. T., Kaspi V. M., Kondratiev V. I., Langer N., Marsh T. R., McLaughlin M. A., Pennucci T. T., Ransom S. M., Stairs I. H., van Leeuwen J., Verbiest J. P. W., Whelan D. G., 2013, Science, 340, 448 (Cited on pages 5, 23 and 27.)

Archibald A. M., Gusinskaia N. V., Hessels J. W. T., Deller A. T., Kaplan D. L., Lorimer D. R., Lynch R. S., Ransom S. M., Stairs I. H., 2018, Nature, 559, 73 (Cited on pages 23, 52 and 150.)

Arzoumanian Z., Chernoff D. F., Cordes J. M., 2002, ApJ, 568, 289 (Cited on page 111.)

Athanasiadis T. M., Berezina M., Antoniadis J., Champion D. J., Cruces M., Spitler L., Kramer M., 2021, MNRAS, 505, 4981 (Cited on page 184.)

Baade W., Zwicky F., 1934, Proceedings of the National Academy of Sciences, 20, 259 (Cited on page 4.)

Backer D., Kulkarni S., Heiles C., Davis M., Goss M., 1982, IAU Circ., 3743, 2 (Cited on page 45.)

Backer D. C., 1970a, Nature, 228, 752 (Cited on page 17.)

Backer D. C., 1970b, Nature, 228, 42 (Cited on page 17.)

Backer D. C., 1975, A&A, 43, 395 (Cited on page 12.)

Bagchi M., Lorimer D. R., Wolfe S., 2013, MNRAS, 432, 1303 (Cited on pages 108, 114 and 173.)

Bailes M., Bates S. D., Bhalerao V., Bhat N. D. R., Burgay M., Burke-Spolaor S., D'Amico N., Johnston S., Keith M. J., Kramer M., Kulkarni S. R., Levin L., Lyne A. G., Milia S., Possenti A., Spitler L., Stappers B., van Straten W., 2011, Science, 333, 1717 (Cited on pages 49 and 50.)

Balakrishnan V., Champion D., Barr E., Kramer M., Sengar R., Bailes M., 2021, Monthly Notices of the Royal Astronomical Society, 505, 1180 (Cited on pages 106 and 121.)

Balakrishnan V., Champion D., Barr E., Kramer M., Venkatraman Krishnan V., Eatough R. P., Sengar R., Bailes M., 2022, MNRAS, 511, 1265 (Cited on page 149.)

Barr E. D., Champion D. J., Kramer M., Eatough R. P., Freire P. C. C., Karuppusamy R., Lee K. J., Verbiest J. P. W., Bassa C. G., Lyne A. G., Stappers B., Lorimer D. R., Klein B., 2013, MNRAS, 435, 2234 (Cited on pages 42 and 46.)

Barsdell B. R., Bailes M., Barnes D. G., Fluke C. J., 2012, MNRAS, 422, 379 (Cited on pages 71 and 177.)

Bates S. D., Bailes M., Barsdell B. R., Bhat N. D. R., Burgay M., Burke-Spolaor S., Champion D. J., Coster P., D'Amico N., Jameson A., Johnston S., Keith M. J., Kramer M., Levin L., Lyne A., Milia S., Ng C., Nietner C., Possenti A., Stappers B., Thornton D., van Straten W., 2012, MNRAS, 427, 1052 (Cited on page 124.)

Bates S. D., Bailes M., Bhat N. D. R., Burgay M., Burke-Spolaor S., D'Amico N., Jameson A., Johnston S., Keith M. J., Kramer M., Levin L., Lyne A., Milia S., Possenti A., Stappers B., van Straten W., 2011, MNRAS, 416, 2455 (Cited on page 49.)

Bates S. D., Lorimer D. R., Rane A., Swiggum J., 2014, MNRAS, 439, 2893 (Cited on pages 106 and 107.)

Bates S. D., Thornton D., Bailes M., Barr E., Bassa C. G., Bhat N. D. R., Burgay M., Burke-Spolaor S., Champion D. J., Flynn C. M. L., Jameson A., Johnston S., Keith M. J., Kramer M., Levin L., Lyne A., Milia S., Ng C., Petroff E., Possenti A., Stappers B. W., van Straten W., Tiburzi C., 2015, MNRAS, 446, 4019 (Cited on page 49.)

Baym G., Pethick C., Pines D., 1969, Nature, 224, 673 (Cited on page 27.)

Bertotti B., Iess L., Tortora P., 2003, Nature, 425, 374 (Cited on page 22.)

Bethapudi S., Desai S., 2018, Astronomy and Computing, 23, 15 (Cited on pages 124 and 183.)

Bhat N. D. R., Cordes J. M., Camilo F., Nice D. J., Lorimer D. R., 2004, ApJ, 605, 759 (Cited on page 12.)

Bhattacharya D., van den Heuvel E. P. J., 1991, Phys. Rep., 203, 1 (Cited on page 16.)

Bhattacharyya B., Roy J., Johnson T. J., Ray P. S., Freire P. C. C., Gupta Y., Bhattacharya D., Kaninghat A., Stappers B. W., Ferrara E. C., Sengupta S., Rathour R. S., Kerr M., Smith D. A., Saz Parkinson P. M., Ransom S. M., Michelson P. F., 2021, ApJ, 910, 160 (Cited on page 85.)

Biggs J. D., 1992, ApJ, 394, 574 (Cited on page 17.)

Bignami G. F., Caraveo P. A., De Luca A., Mereghetti S., 2003, Nature, 423, 725 (Cited on page 9.)

Bishop C. M., 2006, Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag, Berlin, Heidelberg (Cited on page 124.)

Blandford R., Teukolsky S. A., 1976, ApJ, 205, 580 (Cited on pages 20, 82 and 177.)

Bracewell R. N., 2000, The Fourier transform and its applications (Cited on page 40.)

Breton R. P., Kaspi V. M., Kramer M., McLaughlin M. A., Lyutikov M., Ransom S. M., Stairs I. H., Ferdman R. D., Camilo F., Possenti A., 2008, Science, 321, 104 (Cited on pages 17 and 23.)

Breton R. P., Roberts M. S. E., Ransom S. M., Kaspi V. M., Durant M., Bergeron P., Faulkner A. J., 2007, ApJ, 661, 1073 (Cited on page 89.)

Brown W. R., Kilic M., Kosakowski A., Andrews J. J., Heinke C. O., Agüeros M. A., Camilo F., Gianninas A., Hermes J. J., Kenyon S. J., 2020, ApJ, 889, 49 (Cited on page 88.)

Bue B. D., Wagstaff K. L., Rebbapragada U. D., Thompson D. R., Tang B., 2014, in Proc. 2014 Conf. on 'Big Data from Space (BiDS'14) Vol. 206, Astronomical data triage for rapid science return (Cited on page 123.)

Burgay M., Bailes M., Bates S. D., Bhat N. D. R., Burke-Spolaor S., Champion D. J., Coster P., D'Amico N., Johnston S., Keith M. J., Kramer M., Levin L., Lyne A. G., Milia S., Ng C., Possenti A., Stappers B. W., Thornton D., Tiburzi C., van Straten W., Bassa C. G., 2013, MNRAS, 433, 259 (Cited on page 49.)

Burgay M., D'Amico N., Possenti A., Manchester R. N., Lyne A. G., Joshi B. C., McLaughlin M. A., Kramer M., Sarkissian J. M., Camilo F., Kalogera V., Kim C., Lorimer D. R., 2003, Nature, 426, 531 (Cited on pages 17, 23, 150 and 162.)

Burke B. F., Graham-Smith F., 2009, An Introduction to Radio Astronomy (Cited on page 32.)

Burns W. R., Clark B. G., 1969, A&A, 2, 280 (Cited on page 45.)

Cameron A., 2018, PhD thesis, Rheinische Friedrich-Wilhelms-Universität Bonn (Cited on pages 33 and 48.)

Cameron A. D., Champion D. J., Bailes M., Balakrishnan V., Barr E. D., Bassa C. G., Bates S., Bhandari S., Bhat N. D. R., Burgay M., Burke-Spolaor S., Flynn C. M. L., Jameson A., Johnston S., Keith M. J., Kramer M., Levin L., Lyne A. G., Ng C., Petroff E., Possenti A., Smith D. A., Stappers B. W., van Straten W., Tiburzi C., Wu J., 2020, MNRAS, 493, 1063 (Cited on pages 42, 49, 50, 58, 59, 71, 73, 89, 98, 100, 102, 106, 107, 114, 118, 136, 172, 181 and 182.)

Cameron A. D., Champion D. J., Kramer M., Bailes M., Barr E. D., Bassa C. G., Bhandari S., Bhat N. D. R., Burgay M., Burke-Spolaor S., Eatough R. P., Flynn C. M. L., Freire P. C. C., Jameson A., Johnston S., Karuppusamy R., Keith M. J., Levin L., Lorimer D. R., Lyne A. G., McLaughlin M. A., Ng C., Petroff E., Possenti A., Ridolfi A., Stappers B. W., van Straten W., Tauris T. M., Tiburzi C., Wex N., 2018, MNRAS, 475, L57 (Cited on pages 50, 98, 162 and 182.)

Camilo F., Lorimer D. R., Bhat N. D. R., Gotthelf E. V., Halpern J. P., Wang Q. D., Lu F. J., Mirabal N., 2002, ApJ, 574, L71 (Cited on page 14.)

Camilo F., Lorimer D. R., Freire P., Lyne A. G., Manchester R. N., 2000, ApJ, 535, 975 (Cited on pages 21, 41 and 85.)

Camilo F., Ransom S. M., Chatterjee S., Johnston S., Demorest P., 2012, ApJ, 746, 63 (Cited on page 17.)

Camilo F., Ransom S. M., Halpern J. P., Reynolds J., Helfand D. J., Zimmerman N., Sarkissian J., 2006, Nature, 442, 892 (Cited on page 19.)

Camilo F., Ransom S. M., Peñalver J., Karastergiou A., van Kerkwijk M. H., Durant M., Halpern J. P., Reynolds J., Thum C., Helfand D. J., Zimmerman N., Cognard I., 2007, ApJ, 669, 561 (Cited on page 19.)

Cardall C. Y., Prakash M., Lattimer J. M., 2001, ApJ, 554, 322 (Cited on page 5.)

Chandrasekhar S., 1931, ApJ, 74, 81 (Cited on page 5.)

Chattopadhyay D., Stevenson S., Hurley J. R., Bailes M., Broekgaarden F., 2021, MNRAS, 504, 3682 (Cited on pages 53, 59, 172 and 185.)

Chen K., Ruderman M., 1993, ApJ, 402, 264 (Cited on pages 6, 14 and 15.)

Chollet F., et al.„ 2015, Keras, `https://keras.io` (Cited on pages 136 and 147.)

Clausen D., Sigurdsson S., Chernoff D. F., 2014, MNRAS, 442, 207 (Cited on pages 53 and 175.)

Clifton T. R., Lyne A. G., 1986, Nature, 320, 43 (Cited on page 45.)

Colpi M., Mapelli M., Possenti A., 2003, ApJ, 599, 1260 (Cited on page 175.)

Cooley J., Tukey J., 1965, Mathematics of Computation, 19, 297 (Cited on pages 38, 40, 57, 122 and 156.)

Cordes J. M., Freire P. C. C., Lorimer D. R., Camilo F., Champion D. J., Nice D. J., Ramachand ran R., Hessels J. W. T., Vlemmings W., van Leeuwen J., Ransom S. M., Bhat N. D. R., Arzoumanian Z., McLaughlin M. A., Kaspi V. M., Kasian L., Deneva J. S., Reid B., Chatterjee S., Han J. L., Backer D. C., Stairs I. H., Deshpand e A. A., Faucher-Giguère C. A., 2006, ApJ, 637, 446 (Cited on pages 123, 151, 162 and 174.)

Cordes J. M., Lazio T. J. W., 2002, arXiv e-prints, pp astro–ph/0207156 (Cited on pages 10, 74, 75, 98 and 104.)

Dai Z., Yang Z., Yang F., Cohen W. W., Salakhutdinov R., 2017, arXiv e-prints, p. arXiv:1705.09783 (Cited on page 147.)

Dall'Osso S., Granot J., Piran T., 2012, MNRAS, 422, 2878 (Cited on page 19.)

Damour T., Esposito-Farese G., 1992, Classical and Quantum Gravity, 9, 2093 (Cited on page 23.)

Damour T., Esposito-Farese G., 1993, Phys. Rev. Lett., 70, 2220 (Cited on page 52.)

Damour T., Esposito-Farèse G., 1996, Phys. Rev. D, 54, 1474 (Cited on page 52.)

Damour T., Nordtvedt K., 1993, Phys. Rev. D, 48, 3436 (Cited on page 23.)

Damour T., Schaefer G., 1991, Phys. Rev. Lett., 66, 2549 (Cited on page 23.)

Davies J. G., Large M. I., 1970, MNRAS, 149, 301 (Cited on page 43.)

Davies J. G., Large M. I., Pickwick A. C., 1970, Nature, 227, 1123 (Cited on page 43.)

Davies J. G., Lyne A. G., Seiradakis J. H., 1973, Nature Physical Science, 244, 84 (Cited on page 43.)

Davies M. B., Ritter H., King A., 2002, MNRAS, 335, 369 (Cited on page 53.)

Deruelle N., 2011, General Relativity and Gravitation, 43, 3337 (Cited on page 52.)

Devecchi B., Colpi M., Mapelli M., Possenti A., 2007, MNRAS, 380, 691 (Cited on page 175.)

Devine T. R., 2020, Graduate theses, dissertations, and problem reports. 7727, West Virginia University (Cited on page 123.)

Dewdney P. E., Hall P. J., Schilizzi R. T., Lazio T. J. L. W., 2009, Proceedings of the IEEE, 97, 1482 (Cited on page 180.)

Dhurandhar S. V., Vecchio A., 2001, Phys. Rev. D, 63, 122001 (Cited on pages 55, 152, 156, 179 and 201.)

Duncan R. C., Thompson C., 1992, ApJ, 392, L9 (Cited on page 19.)

Dyson F. W., Eddington A. S., Davidson C., 1920, Philosophical Transactions of the Royal Society of London Series A, 220, 291 (Cited on page 22.)

Eatough R., Karuppusamy R., Kramer M., Klein B., Champion D., Kraus A., Keane E., Bassa C., Lyne A., Lazarus P., Verbiest J., Freire P., Brunthaler A., Falcke H., 2013, The Astronomer's Telegram, 5040, 1 (Cited on page 19.)

Eatough R. P., Kramer M., Lyne A. G., Keith M. J., 2013, MNRAS, 431, 292 (Cited on pages 41, 46 and 151.)

Eatough R. P., Molkenthin N., Kramer M., Noutsos A., Keith M. J., Stappers B. W., Lyne A. G., 2010a, MNRAS, 407, 2443 (Cited on page 46.)

Eatough R. P., Molkenthin N., Kramer M., Noutsos A., Keith M. J., Stappers B. W., Lyne A. G., 2010b, MNRAS, 407, 2443 (Cited on pages 123 and 124.)

Ebbens A. T., 2020, European Journal of Physics, 41, 025601 (Cited on page 9.)

Edwards R. T., Bailes M., 2001, ApJ, 547, L37 (Cited on page 87.)

Einstein A., 1915a, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin, pp 844–847 (Cited on page 21.)

Einstein A., 1915b, Sitzungsber. preuss.Akad. Wiss, 47, 831 (Cited on page 22.)

Einstein A., 1916, Naturwissenschaften, 4, 481 (Cited on page 22.)

Espinoza C. M., Lyne A. G., Stappers B. W., Kramer M., 2011, MNRAS, 414, 1679 (Cited on page 8.)

Faucher-Giguère C.-A., Kaspi V. M., 2006, ApJ, 643, 332 (Cited on pages 23, 111, 114 and 116.)

Faulkner A. J., Kramer M., Lyne A. G., Manchester R. N., McLaughlin M. A., Stairs I. H., Hobbs G., Possenti A., Lorimer D. R., D'Amico N., Camilo F., Burgay M., 2005, ApJ, 618, L119 (Cited on pages 46 and 182.)

Faulkner A. J., Stairs I. H., Kramer M., Lyne A. G., Hobbs G., Possenti A., Lorimer D. R., Manchester R. N., McLaughlin M. A., D'Amico N., Camilo F., Burgay M., 2004, MNRAS, 355, 147 (Cited on page 42.)

Fonseca E., Cromartie H. T., Pennucci T. T., Ray P. S., Kirichenko A. Y., Ransom S. M., Demorest P. B., Stairs I. H., Arzoumanian Z., Guillemot L., Parthasarathy A., Kerr M., Cognard I., Baker P. T., Blumer H., Brook P. R., DeCesar M., Dolch T., Dong F. A., Ferrara E. C., Fiore W., Garver-Daniels N., Good D. C., Jennings R., Jones M. L., Kaspi V. M., Lam M. T., Lorimer D. R., Luo J., McEwen A., McKee J. W., McLaughlin M. A., McMann N., Meyers B. W., Naidu A., Ng C., Nice D. J., Pol N., Radovan H. A., Shapiro-Albert B., Tan C. M., Tendulkar S. P., Swiggum J. K., Wahl H. M., Zhu W. W., 2021, The Astrophysical Journal Letters, 915, L12 (Cited on page 5.)

Foreman-Mackey D., 2016, The Journal of Open Source Software, 1, 24 (Cited on pages 163 and 180.)

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306 (Cited on pages 162 and 180.)

Foster R. S., Backer D. C., 1990, ApJ, 361, 300 (Cited on page 27.)

Foster R. S., Fairhead L., Backer D. C., 1991, ApJ, 378, 687 (Cited on page 19.)

Freire P. C., 2000, PhD thesis, - (Cited on page 80.)

Freire P. C., Camilo F., Kramer M., Lorimer D. R., Lyne A. G., Manchester R. N., D'Amico N., 2003, MNRAS, 340, 1359 (Cited on page 87.)

Freire P. C., Kramer M., Lyne A. G., 2001, MNRAS, 322, 885 (Cited on pages 20 and 79.)

Freire P. C. C., Ridolfi A., 2018, MNRAS, 476, 4794 (Cited on page 82.)

Freire P. C. C., Wex N., Esposito-Farèse G., Verbiest J. P. W., Bailes M., Jacoby B. A., Kramer M., Stairs I. H., Antoniadis J., Janssen G. H., 2012, MNRAS, 423, 3328 (Cited on page 23.)

Gold T., 1968, Nature, 218, 731 (Cited on page 4.)

Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y., 2014, in Ghahramani Z., Welling M., Cortes C., Lawrence N. D., Weinberger K. Q., eds, , Advances in Neural Information Processing Systems 27. Curran Associates, Inc., pp 2672–2680 (Cited on pages 127 and 128.)

Goodman J., Weare J., 2010, Communications in Applied Mathematics and Computational Science, 5, 65 (Cited on pages 162 and 163.)

Graham-Smith F., 2003, Reports on Progress in Physics, 66, 173 (Cited on page 6.)

Greenstein J. L., Oke J. B., Shipman H. L., 1971, ApJ, 169, 563 (Cited on page 22.)

Guo P., Duan F., Wang P., Yao Y., Yin Q., Xin X., Li D., Qian L., Wang S., Pan Z., Zhang L., 2019, MNRAS, 490, 5424 (Cited on pages 123, 125, 130 and 183.)

Gupta Y., Gangadhara R. T., 2003, ApJ, 584, 418 (Cited on page 85.)

Han J., 2013, in Kosovichev A. G., de Gouveia Dal Pino E., Yan Y., eds, Solar and Astrophysical Dynamos and Magnetic Activity Vol. 294, Magnetic fields in our Milky Way Galaxy and nearby galaxies. pp 213–224 (Cited on page 27.)

Han J. L., Wang C., Wang P. F., Wang T., Zhou D. J., Sun J.-H., Yan Y., Su W.-Q., Jing W.-C., Chen X., Gao X. Y., Hou L.-G., Xu J., Lee K. J., Wang N., Jiang P., Xu R.-X., Yan J., Gan H.-Q., Guan X., Huang W.-J., Jiang J.-C., Li H., Men Y.-P., Sun C., Wang B.-J., Wang H. G., Wang S.-Q., Xie J.-T., Xu H., Yao R., You X.-P., Yu D. J., Yuan J.-P., Yuen R., Zhang C.-F., Zhu Y., 2021, Research in Astronomy and Astrophysics, 21, 107 (Cited on page 182.)

Harry I. W., Allen B., Sathyaprakash B. S., 2009, Phys. Rev. D, 80, 104014 (Cited on pages 62, 66, 136 and 163.)

Haslam C. G. T., Klein U., Salter C. J., Stoffel H., Wilson W. E., Cleary M. N., Cooke D. J., Thomasson P., 1981, A&A, 100, 209 (Cited on page 48.)

Hastings W. K., 1970, Biometrika, 57, 97 (Cited on page 65.)

He K., Zhang X., Ren S., Sun J., 2015, arXiv e-prints, p. arXiv:1512.03385 (Cited on page 146.)

Hewish A., Bell S. J., Pilkington J. D. H., Scott P. F., Collins R. A., 1968, Nature, 217, 709 (Cited on pages 4, 10 and 43.)

Hills J. G., 1976, MNRAS, 175, 1P (Cited on pages 53 and 175.)

Hobbs G., Faulkner A., Stairs I. H., Camilo F., Manchester R. N., Lyne A. G., Kramer M., D'Amico N., Kaspi V. M., Possenti A., McLaughlin M. A., Lorimer D. R., Burgay M., Joshi B. C., Crawford F., 2004, MNRAS, 352, 1439 (Cited on page 46.)

Hobbs G., Lyne A. G., Kramer M., 2010, MNRAS, 402, 1027 (Cited on page 8.)

Hobbs G., Lyne A. G., Kramer M., Martin C. E., Jordan C., 2004, MNRAS, 353, 1311 (Cited on page 89.)

Hobbs G., Manchester R. N., Dunning A., Jameson A., Roberts P., George D., Green J. A., Tuthill J., Toomey L., Kaczmarek J. F., et al. 2020, Publications of the Astronomical Society of Australia, 37, e012 (Cited on page 74.)

Hotan A. W., van Straten W., Manchester R. N., 2004, PASA, 21, 302 (Cited on page 34.)

Hulse R. A., 1975, PhD thesis, Massachusetts Univ., Amherst. (Cited on page 45.)

Hulse R. A., Taylor J. H., 1974, ApJ, 191, L59 (Cited on page 45.)

Hulse R. A., Taylor J. H., 1975a, ApJ, 201, L55 (Cited on page 45.)

Hulse R. A., Taylor J. H., 1975b, ApJ, 195, L51 (Cited on pages 17, 22, 45, 122 and 150.)

Hunter J. D., 2007, Computing in Science & Engineering, 9, 90 (Cited on pages 147 and 180.)

Ihm C. M., Kalogera V., Belczynski K., 2006, ApJ, 652, 540 (Cited on page 16.)

Ivanova N., Heinke C. O., Rasio F. A., Belczynski K., Fregeau J. M., 2008, MNRAS, 386, 553 (Cited on pages 53 and 175.)

Jackson J. D., 1962, Classical Electrodynamics (Cited on page 8.)

Jaffe A. H., Backer D. C., 2003, ApJ, 583, 616 (Cited on page 27.)

Jenet F. A., Hobbs G. B., Lee K. J., Manchester R. N., 2005, ApJ, 625, L123 (Cited on page 27.)

Johnston H. M., Kulkarni S. R., 1991, ApJ, 368, 504 (Cited on pages 54, 71, 108, 151, 158 and 176.)

Johnston S., Karastergiou A., 2017, MNRAS, 467, 3493 (Cited on page 14.)

Johnston S., Lyne A. G., Manchester R. N., Kniffen D. A., D'Amico N., Lim J., Ashworth M., 1992, MNRAS, 255, 401 (Cited on page 45.)

Johnston S., Manchester R. N., Lyne A. G., Bailes M., Kaspi V. M., Qiao G., D'Amico N., 1992, ApJ, 387, L37 (Cited on page 88.)

Jones D. L., Wagstaff K., Thompson D. R., D'Addario L., Navarro R., Mattmann C., Majid W., Lazio J., Preston R., Rebbapragada U., 2012, in 2012 IEEE Aerospace Conference Big data challenges for large radio arrays. pp 1–6 (Cited on page 123.)

Kaplan D. L., Dai S., Lenc E., Zic A., Swiggum J. K., Murphy T., Anderson C. S., Cameron A. D., Dobie D., Hobbs G., Kaczmarek J. F., Lynch C., Toomey L., 2019, ApJ, 884, 96 (Cited on page 106.)

Karras T., Laine S., Aila T., 2018, arXiv e-prints, p. arXiv:1812.04948 (Cited on page 128.)

Karuppusamy R., Desvignes G., Kramer M., Porayko N., Champion D., Torne P., Stappers B., van der Horst A., Kouveliotou C., O'Connor B., 2020, The Astronomer's Telegram, 13553, 1 (Cited on page 19.)

Kaspi V. M., Beloborodov A. M., 2017, ARA&A, 55, 261 (Cited on page 19.)

Kaspi V. M., Helfand D. J., 2002, in Slane P. O., Gaensler B. M., eds, Neutron Stars in Supernova Remnants Vol. 271 of Astronomical Society of the Pacific Conference Series, Constraining the Birth Events of Neutron Stars. p. 3 (Cited on page 9.)

Kaspi V. M., Lyne A. G., Manchester R. N., Crawford F., Camilo F., Bell J. F., D'Amico N., Stairs I. H., McKay N. P. F., Morris D. J., Possenti A., 2000, ApJ, 543, 321 (Cited on page 46.)

Keane E. F., Ludovici D. A., Eatough R. P., Kramer M., Lyne A. G., McLaughlin M. A., Stappers B. W., 2010, MNRAS, 401, 1057 (Cited on page 46.)

Keith M. J., Jameson A., van Straten W., Bailes M., Johnston S., Kramer M., Possenti A., Bates S. D., Bhat N. D. R., Burgay M., Burke-Spolaor S., D'Amico N., Levin L., McMahon P. L., Milia S., Stappers B. W., 2010, MNRAS, 409, 619 (Cited on pages 26, 46, 47, 49, 106, 107, 122, 136, 170 and 172.)

Keith M. J., Johnston S., Bailes M., Bates S. D., Bhat N. D. R., Burgay M., Burke-Spolaor S., D'Amico N., Jameson A., Kramer M., Levin L., Milia S., Possenti A., Stappers B. W., van Straten W., Parent D., 2012, MNRAS, 419, 1752 (Cited on pages 49 and 89.)

Keith M. J., Kramer M., Lyne A. G., Eatough R. P., Stairs I. H., Possenti A., Camilo F., Manchester R. N., 2009, MNRAS, 393, 623 (Cited on page 46.)

Kingma D. P., Ba J., 2014, arXiv e-prints, p. arXiv:1412.6980 (Cited on page 139.)

Knispel B., 2011, PhD thesis, Albert-Einstein-Institut, Max-Planck-Institut für Gravitationsphysik, D-30167 Hannover, Germany ; Institut für Gravitationsphysik, Leibniz Universität Hannover, D-30167 Hannover, Germany (Cited on pages 55, 57, 62, 65, 68, 71, 151, 154, 156, 164, 174, 177 and 181.)

Knispel B., Eatough R. P., Kim H., Keane E. F., Allen B., Anderson D., Aulbert C., Bock O., Crawford F., Eggenstein H. B., Fehrmann H., Hammer D., Kramer M., Lyne A. G., Machenschalk B., Miller R. B., Papa M. A., Rastawicki D., Sarkissian J., Siemens X., Stappers B. W., 2013, ApJ, 774, 93 (Cited on pages 42, 46, 55, 62, 65, 151, 156, 158, 174 and 181.)

Kocz J., Bailes M., Barnes D., Burke-Spolaor S., Levin L., 2012, MNRAS, 420, 271 (Cited on page 37.)

Kramer et al., 2021, submitted (Cited on page 150.)

Kramer M., 2004, arXiv e-prints, pp astro–ph/0409020 (Cited on pages 114 and 150.)

Kramer M., Backer D. C., Cordes J. M., Lazio T. J. W., Stappers B. W., Johnston S., 2004, New A Rev., 48, 993 (Cited on pages 52 and 53.)

Kramer M., Bell J. F., Manchester R. N., Lyne A. G., Camilo F., Stairs I. H., D'Amico N., Kaspi V. M., Hobbs G., Morris D. J., Crawford F., Possenti A., Joshi B. C., McLaughlin M. A., Lorimer D. R., Faulkner A. J., 2003, MNRAS, 342, 1299 (Cited on page 46.)

Kramer M., Champion D. J., 2013, Classical and Quantum Gravity, 30, 224009 (Cited on page 27.)

Kramer M., Lyne A. G., O'Brien J. T., Jordan C. A., Lorimer D. R., 2006, Science, 312, 549 (Cited on page 17.)

Kramer M., Stairs I. H., Manchester R. N., McLaughlin M. A., Lyne A. G., Ferdman R. D., Burgay M., Lorimer D. R., Possenti A., D'Amico N., Sarkissian J. M., Hobbs G. B., Reynolds J. E., Freire P. C. C., Camilo F., 2006, Science, 314, 97 (Cited on pages 17, 23 and 150.)

Kramer M., Wielebinski R., Jessner A., Gil J. A., Seiradakis J. H., 1994, A & A, 107, 515 (Cited on pages 19 and 85.)

Kraus J. D., 1966, Radio astronomy (Cited on page 33.)

Kruckow M. U., Tauris T. M., Langer N., Kramer M., Izzard R. G., 2018, MNRAS, 481, 1908 (Cited on page 54.)

Lange C., Camilo F., Wex N., Kramer M., Backer D. C., Lyne A. G., Doroshenko O., 2001, MNRAS, 326, 274 (Cited on pages 20, 82 and 177.)

Large M. I., Vaughan A. E., Mills B. Y., 1968, Nature, 220, 340 (Cited on page 4.)

Large M. I., Vaughan A. E., Wielebinski R., 1968, Nature, 220, 753 (Cited on page 43.)

Lattimer J. M., 2012, Annual Review of Nuclear and Particle Science, 62, 485 (Cited on page 27.)

Lattimer J. M., Prakash M., 2001, ApJ, 550, 426 (Cited on page 5.)

Lazarus P., Brazier A., Hessels J. W. T., Karako-Argaman C., Kaspi V. M., Lynch R., Madsen E., Patel C., Ransom S. M., Scholz P., Swiggum J., Zhu W. W., Allen B., Bogdanov S., Camilo F., Cardoso F., Chatterjee S., Cordes J. M., Crawford F., Deneva J. S., Ferdman R., Freire P. C. C., Jenet F. A., Knispel B., Lee K. J., van

Leeuwen J., Lorimer D. R., Lyne A. G., McLaughlin M. A., Siemens X., Spitler L. G., Stairs I. H., Stovall K., Venkataraman A., 2015, ApJ, 812, 81 (Cited on pages 26, 42, 107 and 108.)

LeCun Y., Bottou L., Bengio Y., Haffner P., 1998, Proceedings of the IEEE, 86, 2278 (Cited on page 125.)

Lee K. J., Stovall K., Jenet F. A., Martinez J., Dartez L. P., Mata A., Lunsford G., Cohen S., Biwer C. M., Rohr M., Flanigan J., Walker A., Banaszak S., Allen B., Barr E. D., Bhat N. D. R., Bogdanov S., Brazier A., Camilo F., Champion D. J., Chatterjee S., Cordes J., Crawford F., Deneva J., Desvignes G., Ferdman R. D., Freire P., Hessels J. W. T., Karuppusamy R., Kaspi V. M., Knispel B., Kramer M., Lazarus P., Lynch R., Lyne A., McLaughlin M., Ransom S., Scholz P., Siemens X., Spitler L., Stairs I., Tan M., van Leeuwen J., Zhu W. W., 2013, Monthly Notices of the Royal Astronomical Society, 433, 688 (Cited on page 123.)

Léna P., Rouan D., Lebrun F., Mignard F., Pelat D., 2012, Observational Astrophysics (Cited on page 32.)

LeNail A., 2019, Journal of Open Source Software, 4, 747 (Cited on pages 126 and 127.)

Levin L., Bailes M., Barsdell B. R., Bates S. D., Bhat N. D. R., Burgay M., Burke-Spolaor S., Champion D. J., Coster P., D'Amico N., Jameson A., Johnston S., Keith M. J., Kramer M., Milia S., Ng C., Possenti A., Stappers B., Thornton D., van Straten W., 2013, MNRAS, 434, 1387 (Cited on page 107.)

Levin L., Bailes M., Bates S., Bhat N. D. R., Burgay M., Burke-Spolaor S., D'Amico N., Johnston S., Keith M., Kramer M., Milia S., Possenti A., Rea N., Stappers B., van Straten W., 2010, ApJ, 721, L33 (Cited on pages 19 and 50.)

Liu K., Eatough R. P., Wex N., Kramer M., 2014, MNRAS, 445, 3115 (Cited on page 150.)

Löhmer O., Mitra D., Gupta Y., Kramer M., Ahuja A., 2004, A&A, 425, 569 (Cited on page 13.)

Lomb N. R., 1976, Ap&SS, 39, 447 (Cited on page 80.)

Lorimer D. R., 2008, Living Reviews in Relativity, 11, 8 (Cited on pages 18 and 111.)

Lorimer D. R., Faulkner A. J., Lyne A. G., Manchester R. N., Kramer M., McLaughlin M. A., Hobbs G., Possenti A., Stairs I. H., Camilo F., Burgay M., D'Amico N., Corongiu A., Crawford F., 2006, MNRAS, 372, 777 (Cited on pages 46, 88 and 107.)

Lorimer D. R., Kramer M., 2012, Handbook of Pulsar Astronomy (Cited on pages 7, 8, 9, 10, 19, 21, 38 and 154.)

Lorimer D. R., Lyne A. G., McLaughlin M. A., Kramer M., Pavlov G. G., Chang C., 2012, ApJ, 758, 141 (Cited on page 17.)

Lovelace R. V. E., Sutton J. M., Craft H. D., 1968, IAU Circ., 2113, 1 (Cited on page 4.)

Lyne A., Hobbs G., Kramer M., Stairs I., Stappers B., 2010, Science, 329, 408 (Cited on page 17.)

Lyne A. G., 1992, Philosophical Transactions of the Royal Society of London Series A, 341, 29 (Cited on page 27.)

Lyne A. G., 2008, in Bassa C., Wang Z., Cumming A., Kaspi V. M., eds, 40 Years of Pulsars: Millisecond Pulsars, Magnetars and More Vol. 983 of American Institute of Physics Conference Series, Parkes 20-cm Multibeam Pulsar Surveys. pp 561–566 (Cited on page 45.)

Lyne A. G., Burgay M., Kramer M., Possenti A., Manchester R. N., Camilo F., McLaughlin M. A., Lorimer D. R., D'Amico N., Joshi B. C., Reynolds J., Freire P. C. C., 2004, Science, 303, 1153 (Cited on pages 17, 23 and 150.)

Lyne A. G., Camilo F., Manchester R. N., Bell J. F., Kaspi V. M., D'Amico N., McKay N. P. F., Crawford F., Morris D. J., Sheppard D. C., Stairs I. H., 2000, MNRAS, 312, 698 (Cited on page 46.)

Lyne A. G., Manchester R. N., 1988, MNRAS, 234, 477 (Cited on page 85.)

Lyne A. G., Manchester R. N., Lorimer D. R., Bailes M., D'Amico N., Tauris T. M., Johnston S., Bell J. F., Nicastro L., 1998, MNRAS, 295, 743 (Cited on page 45.)

Lyne A. G., Rickett B. J., 1968, Nature, 219, 1339 (Cited on page 12.)

Lyne A. G., Stappers B. W., Bogdanov S., Ferdman R. D., Freire P. C. C., Kaspi V. M., Knispel B., Lynch R., Allen B., Brazier A., Camilo F., Cardoso F., Chatterjee S., Cordes J. M., Crawford F., Deneva J. S., Hessels J. W. T., Jenet F. A., Lazarus P., van Leeuwen J., Lorimer D. R., Madsen E., McKee J., McLaughlin M. A., Parent E., Patel C., Ransom S. M., Scholz P., Seymour A., Siemens X., Spitler L. G., Stairs I. H., Stovall K., Swiggum J., Wharton R. S., Zhu W. W., Aulbert C., Bock O., Eggenstein H. B., Fehrmann H., Machenschalk B., 2017, ApJ, 834, 137 (Cited on page 88.)

Lyne A. G., Stappers B. W., Freire P. C. C., Hessels J. W. T., Kaspi V. M., Allen B., Bogdanov S., Brazier A., Camilo F., Cardoso F., Chatterjee S., Cordes J. M., Crawford F., Deneva J. S., Ferdman R. D., Jenet F. A., Knispel B., Lazarus P., van Leeuwen J., Lynch R., Madsen E., McLaughlin M. A., Parent E., Patel C., Ransom S. M., Scholz P., Seymour A., Siemens X., Spitler L. G., Stairs I. H., Stovall K., Swiggum J., Wharton R. S., Zhu W. W., 2017, ApJ, 834, 72 (Cited on page 17.)

Lyon R. J., Stappers B. W., Cooper S., Brooke J. M., Knowles J. D., 2016, MNRAS, 459, 1104 (Cited on pages 122, 145 and 184.)

Lyons R., 2011, Understanding Digital Signal Processing. Prentice Hall (Cited on page 34.)

Ma Z., Zhu J., Zhu Y., Xu H., 2019, in International Conference on Data Mining and Big Data Classification of radio galaxy images with semi-supervised learning. pp 191–200 (Cited on page 123.)

Manchester R. N., Hobbs G., Bailes M., Coles W. A., van Straten W., Keith M. J., Shannon R. M., Bhat N. D. R., Brown A., Burke-Spolaor S. G., Champion D. J., Chaudhary A., Edwards R. T., Hampson G., Hotan A. W., Jameson A., Jenet F. A., Kesteven M. J., Khoo J., Kocz J., Maciesiak K., Oslowski S., Ravi V., Reynolds J. R., Sarkissian J. M., Verbiest J. P. W., Wen Z. L., Wilson W. E., Yardley D., Yan W. M., You X. P., 2013, PASA, 30, e017 (Cited on page 27.)

Manchester R. N., Hobbs G. B., Teoh A., Hobbs M., 2005, AJ, 129, 1993 (Cited on pages 15, 23 and 31.)

Manchester R. N., IPTA 2013, Classical and Quantum Gravity, 30, 224010 (Cited on page 28.)

Manchester R. N., Lyne A. G., Camilo F., Bell J. F., Kaspi V. M., D'Amico N., McKay N. P. F., Crawford F., Stairs I. H., Possenti A., Kramer M., Sheppard D. C., 2001, MNRAS, 328, 17 (Cited on pages 17, 41, 45, 46, 123, 151, 162 and 174.)

Manchester R. N., Lyne A. G., D'Amico N., Bailes M., Johnston S., Lorimer D. R., Harrison P. A., Nicastro L., Bell J. F., 1996, MNRAS, 279, 1235 (Cited on page 45.)

Manchester R. N., Lyne A. G., Taylor J. H., Durdin J. M., Large M. I., Little A. G., 1978, MNRAS, 185, 409 (Cited on page 45.)

McLaughlin M. A., 2013, Classical and Quantum Gravity, 30, 224008 (Cited on page 28.)

McLaughlin M. A., Lyne A. G., Lorimer D. R., Kramer M., Faulkner A. J., Manchester R. N., Cordes J. M., Camilo F., Possenti A., Stairs I. H., Hobbs G., D'Amico N., Burgay M., O'Brien J. T., 2006, Nature, 439, 817 (Cited on page 46.)

Messenger C., Prix R., Papa M. A., 2009, Phys. Rev. D, 79, 104017 (Cited on pages 62, 65 and 155.)

Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E., 1953, The Journal of Chemical Physics, 21, 1087 (Cited on page 65.)

Meurer A., Smith C. P., Paprocki M., Čertík O., Kirpichev S. B., Rocklin M., Kumar A., Ivanov S., Moore J. K., Singh S., Rathnayake T., Vig S., Granger B. E., Muller R. P., Bonazzi F., Gupta H., Vats S., Johansson F., Pedregosa F., Curry M. J., Terrel A. R., Roučka v., Saboo A., Fernando I., Kulal S., Cimrman R., Scopatz A., 2017, PeerJ Computer Science, 3, e103 (Cited on page 180.)

Mickaliger M. B., 2013, PhD thesis, West Virginia University (Cited on page 76.)

Middleditch J., Deich W., Kulkarni S., 1993, in van Riper K. A., Epstein R. I., Ho C., eds, Isolated Pulsars Searching for Millisecond Pulsars. p. 372 (Cited on pages 40 and 164.)

Miller M. C., Lamb F. K., Dittmann A. J., Bogdanov S., Arzoumanian Z., Gendreau K. C., Guillot S., Harding A. K., Ho W. C. G., Lattimer J. M., Ludlam R. M., Mahmoodifar S., Morsink S. M., Ray P. S., Strohmayer T. E., Wood K. S., Enoto T., Foster R., Okajima T., Prigozhin G., Soong Y., 2019, ApJ, 887, L24 (Cited on page 5.)

Morello V., Barr E. D., Cooper S., Bailes M., Bates S., Bhat N. D. R., Burgay M., Burke-Spolaor S., Cameron A. D., Champion D. J., Eatough R. P., Flynn C. M. L., Jameson A., Johnston S., Keith M. J., Keane E. F., Kramer M., Levin L., Ng C., Petroff E., Possenti A., Stappers B. W., van Straten W., Tiburzi C., 2019, MNRAS, 483, 3673 (Cited on page 106.)

Morello V., Barr E. D., Stappers B. W., Keane E. F., Lyne A. G., 2020, MNRAS, 497, 4654 (Cited on pages 26, 38 and 98.)

Morris D. J., Hobbs G., Lyne A. G., Stairs I. H., Camilo F., Manchester R. N., Possenti A., Bell J. F., Kaspi V. M., Amico N. D., McKay N. P. F., Crawford F., Kramer M., 2002, MNRAS, 335, 275 (Cited on pages 46 and 73.)

Mustafa M., Bard D., Bhimji W., Lukić Z., Al-Rfou R., Kratochvil J. M., 2019, Computational Astrophysics and Cosmology, 6, 1 (Cited on page 130.)

Narayan R., Piran T., Shemi A., 1991, ApJ, 379, L17 (Cited on page 54.)

Ng C., Bailes M., Bates S. D., Bhat N. D. R., Burgay M., Burke-Spolaor S., Champion D. J., Coster P., Johnston S., Keith M. J., Kramer M., Levin L., Petroff E., Possenti A., Stappers B. W., van Straten W., Thornton D., Tiburzi C., Bassa C. G., Freire P. C. C., Guillemot L., Lyne A. G., Tauris T. M., Shannon R. M., Wex N., 2014, MNRAS, 439, 1865 (Cited on page 49.)

Ng C., Champion D. J., Bailes M., Barr E. D., Bates S. D., Bhat N. D. R., Burgay M., Burke-Spolaor S., Flynn C. M. L., Jameson A., Johnston S., Keith M. J., Kramer M., Levin L., Petroff E., Possenti A., Stappers B. W., van Straten W., Tiburzi C., Eatough R. P., Lyne A. G., 2015, MNRAS, 450, 2922 (Cited on pages 42, 49, 50, 54, 58, 71, 73, 98, 99, 100, 102, 106, 107, 108, 110, 112, 113, 114, 118, 136, 151, 172, 181, 190, 191, 192, 193 and 196.)

Ng C., Kruckow M. U., Tauris T. M., Lyne A. G., Freire P. C. C., Ridolfi A., Caiazzo I., Heyl J., Kramer M., Cameron A. D., Champion D. J., Stappers B., 2018, MNRAS, 476, 4315 (Cited on page 50.)

Ng C. W. Y., 2014, PhD thesis, Rheinischen Friedrich-Wilhelms-Universität, Bonn (Cited on pages 14, 48, 49 and 73.)

Nice D. J., Fruchter A. S., Taylor J. H., 1995, ApJ, 449, 156 (Cited on page 45.)

Nieder L., Allen B., Clark C. J., Pletsch H. J., 2020, ApJ, 901, 156 (Cited on pages 155, 179, 186 and 201.)

Nieder L., Clark C. J., Kandel D., Romani R. W., Bassa C. G., Allen B., Ashok A., Cognard I., Fehrmann H., Freire P., Karuppusamy R., Kramer M., Li D., Machenschalk B., Pan Z., Papa M. A., Ransom S. M., Ray P. S., Roy J., Wang P., Wu J., Aulbert C., Barr E. D., Beheshtipour B., Behnke O., Bhattacharyya B., Breton R. P., Camilo F., Choquet C., Dhillon V. S., Ferrara E. C., Guillemot L., Hessels J. W. T., Kerr M., Kwang S. A., Marsh T. R., Mickaliger M. B., Pleunis Z., Pletsch H. J., Roberts M. S. E., Sanpa-arsa S., Steltner B., 2020, ApJ, 902, L46 (Cited on page 179.)

NVIDIA Vingelmann P., Fitzek F. H., , 2020, CUDA, release: 10.2.89 (Cited on page 180.)

Olausen S. A., Kaspi V. M., 2014, ApJS, 212, 6 (Cited on pages 15 and 19.)

Oliphant T. E., 2006, in , Vol. 1, A guide to NumPy. Trelgol Publishing USA (Cited on page 147.)

Oppenheimer J. R., Volkoff G. M., 1939, Physical Review, 55, 374 (Cited on page 5.)

Owen B. J., 1996, Phys. Rev. D, 53, 6749 (Cited on pages 56, 62 and 153.)

Owen B. J., Sathyaprakash B. S., 1999, Phys. Rev. D, 60, 022002 (Cited on pages 56, 62 and 153.)

Özel F., Freire P., 2016, ARA&A, 54, 401 (Cited on page 5.)

Pacini F., 1967, Nature, 216, 567 (Cited on page 4.)

Paczynski B., Trimble V., 1979, in Burton W. B., ed., The Large-Scale Characteristics of the Galaxy Vol. 84, Galactic Center Pulsar as a Test of Black Hole Existence and Properties. p. 401 (Cited on page 52.)

Padmanabh et al., in prep, in preparation (Cited on pages 73, 75, 76, 105 and 182.)

Pedregosa F., Varoquaux G., Gramfort A., Michel V., Thirion B., Grisel O., Blondel M., Prettenhofer P., Weiss R., Dubourg V., Vanderplas J., Passos A., Cournapeau D., Brucher M., Perrot M., Duchesnay E., 2011, Journal of Machine Learning Research, 12, 2825 (Cited on page 147.)

Pence W. D., Chiappetti L., Page C. G., Shaw R. A., Stobie E., 2010, A&A, 524, A42 (Cited on page 35.)

Perera B. B. P., Barr E. D., Mickaliger M. B., Lyne A. G., Lorimer D. R., Stappers B. W., Eatough R. P., Kramer M., Ng C., Spiewak R., Bailes M., Champion D. J., Morello V., Possenti A., 2019, MNRAS, 487, 1025 (Cited on pages 76 and 82.)

Pfahl E., Podsiadlowski P., Rappaport S., 2005, ApJ, 628, 343 (Cited on pages 53, 54 and 173.)

Phinney E. S., Kulkarni S. R., 1994, ARA&A, 32, 591 (Cited on pages 16 and 87.)

Pound R. V., Rebka G. A., 1959, Phys. Rev. Lett., 3, 439 (Cited on page 22.)

Prager B. J., Ransom S. M., Freire P. C. C., Hessels J. W. T., Stairs I. H., Arras P., Cadelano M., 2017, ApJ, 845, 148 (Cited on page 175.)

Press W. H., Flannery B. P., Teukolsky S. A., Vetterling W. T., 1992, Numerical Recipes in FORTRAN 77: The Art of Scientific Computing, 2 edn. Cambridge University Press (Cited on page 80.)

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical recipes in C. The art of scientific computing (Cited on page 40.)

Prix R., 2007, Classical and Quantum Gravity, 24, S481 (Cited on page 62.)

Radford A., Metz L., Chintala S., 2015, arXiv e-prints, p. arXiv:1511.06434 (Cited on page 130.)

Rankin J. M., 1983, ApJ, 274, 333 (Cited on page 85.)

Ransom S., , 2011, PRESTO: PulsaR Exploration and Search TOolkit (Cited on page 136.)

Ransom S. M., 2001, PhD thesis, Harvard University (Cited on page 42.)

Ransom S. M., Cordes J. M., Eikenberry S. S., 2003, ApJ, 589, 911 (Cited on pages 42, 54 and 151.)

Ransom S. M., Eikenberry S. S., Middleditch J., 2002, AJ, 124, 1788 (Cited on pages 40, 41, 42, 54, 68, 158 and 164.)

Reed S. E., Akata Z., Yan X., Logeswaran L., Schiele B., Lee H., 2016, ArXiv, abs/1605.05396 (Cited on page 128.)

Riley T. E., Watts A. L., Ray P. S., Bogdanov S., Guillot S., Morsink S. M., Bilous A. V., Arzoumanian Z., Choudhury D., Deneva J. S., Gendreau K. C., Harding A. K., Ho W. C. G., Lattimer J. M., Loewenstein M., Ludlam R. M., Markwardt C. B., Okajima T., Prescod-Weinstein C., Remillard R. A., Wolff M. T., Fonseca E., Cromartie H. T., Kerr M., Pennucci T. T., Parthasarathy A., Ransom S., Stairs I., Guillemot L., Cognard I., 2021, arXiv e-prints, p. arXiv:2105.06980 (Cited on page 5.)

Ritchings R. T., 1976, MNRAS, 176, 249 (Cited on page 17.)

Rohlfs K., Wilson T. L., 2000, Tools of radio astronomy (Cited on page 32.)

Roy A. E., 1988, Journal of the British Astronomical Association, 98, 374 (Cited on page 79.)

Ruderman M., 1976, ApJ, 203, 213 (Cited on page 27.)

Rumelhart D. E., Hinton G. E., Williams R. J., 1986, Nature, 323, 533 (Cited on page 125.)

Salimans T., Goodfellow I., Zaremba W., Cheung V., Radford A., Chen X., Chen X., 2016, in Lee D. D., Sugiyama M., Luxburg U. V., Guyon I., Garnett R., eds, , Advances in Neural Information Processing Systems 29. Curran Associates, Inc., pp 2234–2242 (Cited on pages 131, 138 and 146.)

Scargle J. D., 1982, ApJ, 263, 835 (Cited on page 80.)

Schawinski K., Zhang C., Zhang H., Fowler L., Santhanam G. K., 2017, MNRAS, 467, L110 (Cited on page 130.)

Scheuer P. A. G., 1968, Nature, 218, 920 (Cited on page 12.)

Segelstein D. J., 1986, PhD thesis, Princeton Univ., NJ. (Cited on page 45.)

Sengar et al., prep, in preparation (Cited on pages 73, 75, 98, 100, 104, 106, 107, 108, 112, 190, 192 and 196.)

Sesana A., 2013, Brazilian Journal of Physics, 43, 314 (Cited on page 27.)

Shannon R. M., Ravi V., Lentati L. T., Lasky P. D., Hobbs G., Kerr M., Manchester R. N., Coles W. A., Levin Y., Bailes M., Bhat N. D. R., Burke-Spolaor S., Dai S., Keith M. J., Osłowski S., Reardon D. J., van Straten W., Toomey L., Wang J. B., Wen L., Wyithe J. S. B., Zhu X. J., 2015, Science, 349, 1522 (Cited on page 28.)

Shao Y., Li X.-D., 2018, MNRAS, 477, L128 (Cited on pages 151, 172 and 185.)

Shapiro I. I., 1964, Phys. Rev. Lett., 13, 789 (Cited on page 22.)

Shapiro I. I., Ash M. E., Ingalls R. P., Smith W. B., Campbell D. B., Dyce R. B., Jurgens R. F., Pettengill G. H., 1971, Phys. Rev. Lett., 26, 1132 (Cited on page 22.)

Shibazaki N., Murakami T., Shaham J., Nomoto K., 1989, Nature, 342, 656 (Cited on page 16.)

Sigurdsson S., Phinney E. S., 1995, ApJS, 99, 609 (Cited on pages 53 and 175.)

Simonyan K., Zisserman A., 2014, arXiv e-prints, p. arXiv:1409.1556 (Cited on page 146.)

Sipior M. S., Portegies Zwart S., Nelemans G., 2004, MNRAS, 354, L49 (Cited on pages 53 and 173.)

Smith K. M., 2016, arXiv e-prints, p. arXiv:1610.06831 (Cited on page 151.)

Sofue Y., 2013, PASJ, 65, 118 (Cited on page 82.)

Srostlik Z., Rankin J. M., 2005, MNRAS, 362, 1121 (Cited on page 85.)

Staelin D. H., 1969, IEEE Proceedings, 57, 724 (Cited on pages 38 and 122.)

Staelin D. H., Reifenstein Edward C. I., 1968, Science, 162, 1481 (Cited on pages 4 and 14.)

Stairs I. H., Faulkner A. J., Lyne A. G., Kramer M., Lorimer D. R., McLaughlin M. A., Manchester R. N., Hobbs G. B., Camilo F., Possenti A., Burgay M., D'Amico N., Freire P. C., Gregory P. C., 2005, ApJ, 632, 1060 (Cited on page 23.)

Stairs I. H., Manchester R. N., Lyne A. G., Kaspi V. M., Camilo F., Bell J. F., D'Amico N., Kramer M., Crawford F., Morris D. J., Possenti A., McKay N. P. F., Lumsden S. L., Tacconi-Garman L. E., Cannon R. D., Hambly N. C., Wood P. R., 2001, MNRAS, 325, 979 (Cited on page 88.)

Staveley-Smith L., Wilson W. E., Bird T. S., Disney M. J., Ekers R. D., Freeman K. C., Haynes R. F., Sinclair M. W., Vaile R. A., Webster R. L., Wright A. E., 1996, PASA, 13, 243 (Cited on page 45.)

Stokes G. H., Taylor J. H., Welsberg J. M., Dewey R. J., 1985, Nature, 317, 787 (Cited on page 45.)

Stovall K., Lynch R. S., Ransom S. M., Archibald A. M., Banaszak S., Biwer C. M., Boyles J., Dartez L. P., Day D., Ford A. J., Flanigan J., Garcia A., Hessels J. W. T., Hinojosa J., Jenet F. A., Kaplan D. L., Karako-Argaman C., Kaspi V. M., Kondratiev V. I., Leake S., Lorimer D. R., Lunsford G., Martinez J. G., Mata A., McLaughlin M. A., Roberts M. S. E., Rohr M. D., Siemens X., Stairs I. H., van Leeuwen J., Walker A. N., Wells B. L., 2014, ApJ, 791, 67 (Cited on page 123.)

Sturrock P. A., 1971, ApJ, 164, 529 (Cited on page 6.)

Sutton J. M., 1971, MNRAS, 155, 51 (Cited on page 12.)

Szegedy C., Vanhoucke V., Ioffe S., Shlens J., Wojna Z., 2015, arXiv e-prints, p. arXiv:1512.00567 (Cited on page 146.)

Taff L. G., 1985, Celestial mechanics : a computational guide for the practitioner (Cited on page 201.)

Tauris T. M., Langer N., Kramer M., 2011, MNRAS, 416, 2130 (Cited on page 16.)

Tauris T. M., Langer N., Kramer M., 2012, MNRAS, 425, 1601 (Cited on page 9.)

Tauris T. M., Sennels T., 2000, A&A, 355, 236 (Cited on page 53.)

Tauris T. M., van den Heuvel E. P. J., 2006, Formation and evolution of compact stellar X-ray sources. pp 623–665 (Cited on page 16.)

Taylor J. H., Fowler L. A., McCulloch P. M., 1979, Nature, 277, 437 (Cited on page 22.)

Taylor J. H., Huguenin G. R., 1969, Nature, 221, 816 (Cited on pages 40 and 45.)

Taylor J. H., Hulse R. A., Fowler L. A., Gullahorn G. E., Rankin J. M., 1976, ApJ, 206, L53 (Cited on page 22.)

Taylor J. H., Weisberg J. M., 1982, ApJ, 253, 908 (Cited on pages 22 and 150.)

Turtle A. J., Vaughan A. E., 1968, Nature, 219, 689 (Cited on page 43.)

van der Klis M., 1988, in Timing Neutron Stars Vol. 262, Fourier Techniques in X-ray Timing. pp 27–70 (Cited on page 40.)

van Kerkwijk M. H., 1996, in Johnston S., Walker M. A., Bailes M., eds, IAU Colloq. 160: Pulsars: Problems and Progress Vol. 105 of Astronomical Society of the Pacific Conference Series, The White Dwarf Companions of Recycled Pulsars. p. 489 (Cited on page 16.)

van Straten W., Bailes M., 2011, PASA, 28, 1 (Cited on page 80.)

van Straten W., Demorest P., Oslowski S., 2012, Astronomical Research and Technology, 9, 237 (Cited on page 82.)

Vaughan A. E., Large M. I., 1970, Nature, 225, 167 (Cited on page 43.)

Vaughan A. E., Large M. I., 1972, MNRAS, 156, 27P (Cited on page 43.)

Venkatraman Krishnan V., Bailes M., van Straten W., Wex N., Freire P. C. C., Keane E. F., Tauris T. M., Rosado P. A., Bhat N. D. R., Flynn C., Jameson A., Osłowski S., 2020, Science, 367, 577 (Cited on pages 46 and 122.)

Verbiest J. P. W., Bailes M., Coles W. A., Hobbs G. B., van Straten W., Champion D. J., Jenet F. A., Manchester R. N., Bhat N. D. R., Sarkissian J. M., Yardley D., Burke-Spolaor S., Hotan A. W., You X. P., 2009, MNRAS, 400, 951 (Cited on pages 53 and 173.)

Voisin G., Cognard I., Freire P. C. C., Wex N., Guillemot L., Desvignes G., Kramer M., Theureau G., 2020, A&A, 638, A24 (Cited on pages 23, 52, 122 and 150.)

Voss R., Tauris T. M., 2003, MNRAS, 342, 1169 (Cited on pages 53, 151 and 172.)

Wang N., Manchester R. N., Johnston S., 2007, MNRAS, 377, 1383 (Cited on page 17.)

Waskom M., Botvinnik O., O'Kane D., Hobson P., Lukauskas S., Gemperline D. C., Augspurger T., Halchenko Y., Cole J. B., Warmenhoven J., de Ruiter J., Pye C., Hoyer S., Vanderplas J., Villalba S., Kunter G., Quintero E., Bachant P., Martin M., Meyer K., Miles A., Ram Y., Yarkoni T., Williams M. L., Evans C., Fitzgerald C., Brian Fonnesbeck C., Lee A., Qalieh A., 2017, mwaskom/seaborn: v0.8.1 (September 2017) (Cited on page 147.)

Weisberg J. M., Huang Y., 2016, ApJ, 829, 55 (Cited on pages 22, 24 and 150.)

Weisberg J. M., Taylor J. H., 1984, Phys. Rev. Lett., 52, 1348 (Cited on page 150.)

Wen Z. G., Wang N., Yuan J. P., Yan W. M., Manchester R. N., Yuen R., Gajjar V., 2016, A&A, 592, A127 (Cited on page 131.)

Wex N., 2014, arXiv e-prints, p. arXiv:1402.5594 (Cited on page 22.)

Wex N., Kopeikin S. M., 1999, ApJ, 514, 388 (Cited on pages 52 and 150.)

Wielebinski R., Vaughan A. E., Large M. I., 1969, Nature, 221, 47 (Cited on page 43.)

Will C. M., 2014, Living Reviews in Relativity, 17, 4 (Cited on page 22.)

Wolszczan A., Frail D. A., 1992, Nature, 355, 145 (Cited on page 122.)

Woltjer L., 1964, ApJ, 140, 1309 (Cited on page 5.)

Wongphechauxsorn et al., in prep, in preparation (Cited on pages 73, 98, 104 and 105.)

Xu J.-W., Zhang X.-Z., Han J.-L., 2005, Chinese J. Astron. Astrophys., 5, 165 (Cited on page 14.)

Xu Y. H., Lee K. J., Hao L. F., Wang H. G., Liu Z. Y., Yue Y. L., Yuan J. P., Li Z. X., Wang M., Dong J., Tan J. J., Chen W., Bai J. M., 2018, MNRAS, 476, 5579 (Cited on page 131.)

Yan W. M., Manchester R. N., van Straten W., Reynolds J. E., Hobbs G., Wang N., Bailes M., Bhat N. D. R., Burke-Spolaor S., Champion D. J., Coles W. A., Hotan A. W., Khoo J., Oslowski S., Sarkissian J. M., Verbiest J. P. W., Yardley D. R. B., 2011, MNRAS, 414, 2087 (Cited on page 85.)

Yao J. M., Manchester R. N., Wang N., 2017, ApJ, 835, 29 (Cited on pages 10, 38, 74, 75, 98 and 104.)

Young M. D., Manchester R. N., Johnston S., 1999, Nature, 400, 848 (Cited on page 14.)

Yungelson L., Portegies Zwart S. F., 1998, arXiv e-prints, pp astro–ph/9801127 (Cited on pages 53, 151 and 172.)

Zackay B., 2017, in American Astronomical Society Meeting Abstracts #229 Vol. 229 of American Astronomical Society Meeting Abstracts, Algorithms for searching Fast radio bursts and pulsars in tight binary systems.. p. 330.06 (Cited on page 186.)

Zhu W., et al., 2014, ApJ, 781, 117 (Cited on pages 123, 124, 125, 132 and 142.)

Zingales T., Waldmann I. P., 2018, AJ, 156, 268 (Cited on page 130.)

List of Publications

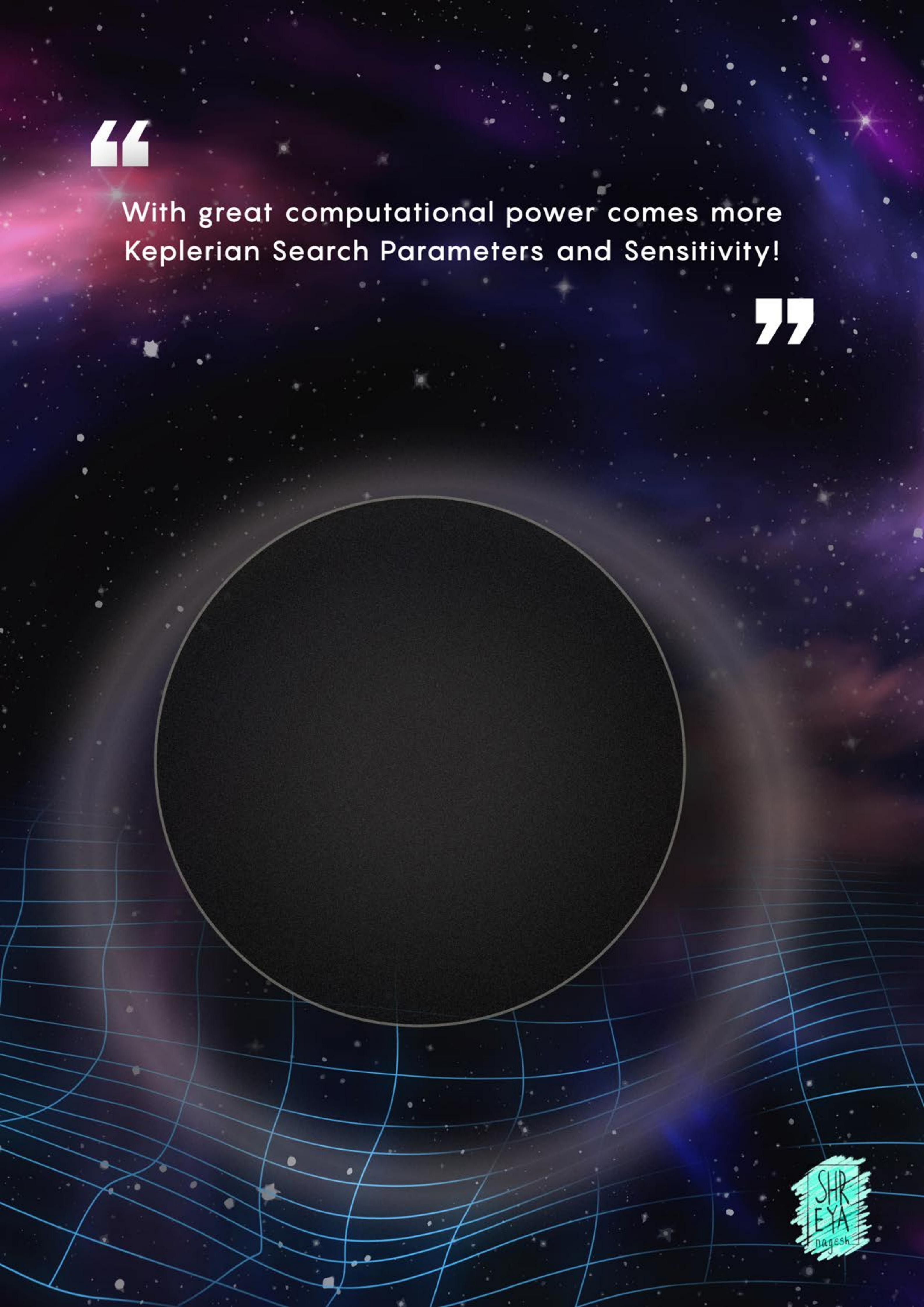
Publications accounting for major parts of the thesis

Refereed:
Pulsar Candidate Identification Using Semi-Supervised Generative Adversarial Networks.
Balakrishnan, V.; Champion, D. J.; Barr, E., Kramer, M., Sengar,R.,Bailes, M.
Monthly Notices of the Royal Astronomical Society, Volume 505, Issue 1, July 2021, Pages 1180–1194.
DOI: `https://doi.org/10.1093/mnras/stab1308`
arXiv: **2010.07457v2**

Coherent Search for Binary Pulsars across all Five Keplerian Parameters in Radio Observations using the template-bank algorithm
Balakrishnan, V.; Champion, D. J.; Barr, E., Kramer, M., Krishnan, V., Eatough, R., Sengar,R. and Bailes,M.
Monthly Notices of the Royal Astronomical Society, Volume 511, Issue 1, March 2022, Pages 1265–1284.
DOI: `https://doi.org/10.1093/mnras/stab3746`
arXiv: **2112.11991**

Other Publications

The High Time Resolution Universe Pulsar Survey - XVI. Discovery and timing of 40 pulsars from the southern Galactic plane.
Cameron, A.; D.; Champion, D. J.; Bailes, M.; **Balakrishnan, V.** et al. 2020
Monthly Notices of the Royal Astronomical Society, Volume 493, Issue 1, March 2020, Pages 1063–1087.
DOI: `https://doi.org/10.1093/mnras/staa039`
arXiv: **2001.01823**



> “With great computational power comes more Keplerian Search Parameters and Sensitivity!”