A Search for Long-Period Pulsars with a Fast-Folding Algorithm in the PALFA Survey

Emilie Parent

A thesis submitted in partial fulfillment for
the degree of Master of Sciences in the

Department of Physics
McGill University, Montreal
Quebec, Canada

Abstract

The PALFA survey, the most sensitive search for radio pulsars, is ongoing at the Arecibo Observatory in Puerto Rico. The survey has now discovered more than 170 pulsars in the Galactic Plane, and the vast majority of these pulsars have periods shorter than 2 seconds. One reason that pulsar surveys may miss long-period radio pulsars is the strong effect of red noise at low modulation frequencies which reduces the sensitivity in this regime, notably for frequencies smaller than 0.1 Hz. It is possible to address this reduction in sensitivity by using a Fast-Folding Algorithm (FFA), a time-domain search technique that has the advantage of coherently summing all harmonics of a given frequency. The FFA, first introduced by Staelin (1969), has been implemented by scientists at Berkeley to look for transients in Kepler data. We have adapted this time-domain search for radio pulsar searching and applied it to PALFA observations. A basic description of the FFA and specific aspects of applying such an algorithm on pulsar searches will be discussed in this MSc thesis. A sensitivity analysis of the algorithm, conducted using synthetic pulsar signals injected in real observational data, will also be presented. This study shows that the FFA recovers significantly the loss in the survey sensitivity at low modulation frequencies, implying that the PALFA survey should discover more long-period radio pulsars in the future.

Résumé

PALFA est le relevé astronomique de pulsars radio le plus sensible au monde. Ce relevé utilise le William E. Gordon Telescope, un radio télescope de 305-m de diamètre à l'Observatoire Arecibo, à Puerto Rico. PALFA a maintenant découvert plus de 170 pulsars dans le disque Galactique. La grande majorité de ces pulsars ont des périodes de rotations inférieures à 2 secondes. Une des raisons pourquoi les relevés astronomiques détectent peu de pulsars ayant de lentes rotations est l'impact nuisible du bruit rouge sur les basses fréquences de modulation, ce qui réduit la sensibilité des relevés dans ce régime (spécialement pour les fréquences de modulation plus petites que 0.1 Hz). Il est possible de (partiellement) remédier à ce problème en utilisant un Fast-Folding Algorithm (FFA), une technique de recherche de signaux périodiques dans le domaine temporel dont l'avantage réside en sa sommation cohérente des composantes harmoniques d'une certaine fréquence fondamentale. Cet algorithme, conçu par Staelin (1969), a été utilisé par des scientifiques à Berkeley dans le but de chercher des transits d'exoplanètes dans les données collectées par le telescope Kepler. Nous avons adapté le programme développé par ces scientifiques pour que ce dernier puisse détecter des signaux de pulsars avec les données recueillies par PALFA. Une description du FFA ainsi que les aspects spécifiques de son application dans notre relevé astronomique seront abordés dans ce mémoire. Une analyse de la performance du FFA, qui a été produite via l'injection de signaux synthétiques de pulsars dans de vrais données de PALFA, sera également présentée. Cette étude démontre qu'en utilisant cet algorithme, notre relevé astronomique est plus sensible aux signaux ayant de longues périodes, ce qui implique que PALFA devrait découvrir plus de pulsars ayant de lentes rotations dans le futur.

Acknowledgments

Firstly, I would like to deeply thank my supervisor, Professor Victoria Kaspi, for her kindness, her remarkable patience and her amazing guidance throughout the project. Doing research under Vicky's supervision has been an incredible experience. I owe special thanks to Chitrang Patel, Paul Scholz and Scott Ransom for their helpful support, especially with regards to the programming aspect of the project. I would also like to extend my gratitude to the PALFA pipeline working group for great explanations and advice. I thank every member of Vicky's research group for their enlightening conversations and their general joviality, which has made my time at McGill unique and enriching. Lastly, I thank my mother and friends for the encouragement they have given me over the course of writing this MSc thesis.

Preface & Contribution from Authors

Several members of the PALFA pipeline working group contributed to the work presented in this MSc thesis: Chitrang Patel, Paul Scholz, Scott Ransom, Victoria Kaspi, Maura McLaughlin, Weiwei Zhu, Adam Brazier, Michelle Boyce, Patrick Lazarus and Erik Madsen. A modified version of the work presented here will eventually be submitted for publication in the Astrophysical Journal.

The contributions of the members are as follows: the part of the FFA program that wraps the data and folds it at the specified period is taken from a github FFA package designed by Erik Petigura (Hubble Fellow at the California Institute of Technology). I have developed the other components of the FFA program, which include the design of the significance metric used to evaluate pulse profiles, the detrending and the dynamic downsampling of the time series, the generation of periodograms for various pulse duty-cycles and the candidate selection. The sifting program, which produces the final list of potential candidates generated by the search, is a slightly modified version of `PRESTO`'s `sifting` program developed by Scott Ransom. Chitrang Patel, helped by Paul Scholz, Scott Ransom and myself, has incorporated the FFA program in the PALFA analysis pipeline, and Chitrang is now responsible for operating the pipeline on the Guillimin supercomputer. Victoria Kaspi coordinates the group and ensures that the survey obtains observational data with the telescope at the Arecibo Observatory, as well as computation time with McGill's Guillimin supercomputer.Data exchanges between the telescope, the PALFA survey database (located at the Cornell University centre for Advanced Computing) and Guillimin are supervised mostly by Adam Brazier and Chitrang Patel.

Adam Brazier also adapted the database so that it can distinguish FFA periodicity candidates from FFT candidates, and Michelle Boyce is in charge of the CyberSKA tool for visualizing data.

The program used for the sensitivity analysis of the Fast-Folding Algorithm is a modified version of software developed by Patrick Lazarus to conduct a sensitivity analysis of the PALFA survey. This work was reported in Lazarus *et al.*, (2015). Parts of the results from that analysis are presented in this MSc thesis.

Contents

List of Figures

Chapter 1

Introduction

1.1 Neutron Stars

1.1.1 Discovery of Neutron Stars

Following the discovery of the neutron by James Chadwick in 1932 (Chadwick, 1932), astronomers Walter Baade and Fritz Zwicky (Baade & Zwicky, 1934) hypothesized the existence of massive stars ($1.4\ \mathrm{M}_{\odot} \lesssim \mathrm{M} \lesssim 3\ \mathrm{M}_{\odot}$, Tolman, 1939, Oppenheimer & Volkoff, 1939) mostly made of densely packed neutrons: neutron stars. These stars are supported by neutron degeneracy pressure and were thought to be produced via the core-collapse supernovae of massive stars as they reach their evolutionary endpoint (Baade & Zwicky, 1934). At the time, the observation of such objects seemed to be impossible since neutron stars should have radii of only a few kilometers because of their great surface gravitational field, as predicted by Oppenheimer & Volkoff (1939). Calculations from Hoyle *et al.* (1964) also suggested the possible existence of a neutron star with a magnetic field of $\sim 10^{10}$ G linked to the X-ray source in the centre of the Crab nebula (Giacconi *et al.*, 1962, Pacini, 1967). It was more than 30 years after Baade and Zwicky proposed their theory that Jocelyn Bell Burnell and her supervisor, Antony Hewish (Hewish *et al.*, 1968), unexpectedly observed the first "pulsar" (PULSating stAR) at Cambridge University: a rotating neutron star emitting beams of radio light. The signal they found was coming from the constellation of Vulpecula, $\sim$700 parsec away from us, and it appeared as periodic pulses separated by 1.33 sec. At first, it was not clear that this signal was emitted by a neutron star, as radio emission from neutron stars was not predicted.

Even if the mechanism behind these radio emissions remains a great mystery in astrophysics, we do know that as the star rotates, periodic pulses are observed if at least one of the two radio beams of the pulsars (which are thought to originate from near the magnetic poles of the star) sweeps our line of sight. Measurements of spin frequencies, which have been found to be as large as 716 Hz (Hessels *et al.*, 2006), suggest that pulsars have very stable rotations. Pulsars will typically complete a few rotations within a second, and the longest spin period we have observed is 8.51 sec (Young *et al.*, 1999). Since their discovery, multiple surveys conducted with large radio telescopes have identified more than 2700 pulsars in our Galaxy.

In the following subsections, different aspects regarding the physics and the observation of neutron stars are presented.

1.1.2 Formation of Neutron Stars

There are two major classes of supernovae: core-collapse (type II, Ib and Ic) and type Ia supernovae. Only the first category will be discussed in this thesis, as those are the kind of explosions that give birth to neutron stars.

By the end of their lives on the main-sequence, stars with masses greater than about 8 solar masses build iron-rich cores as a result of nuclear fusion of lighter elements in their cores, which is an exothermic process. The fusion of elements heavier than iron will result in energy being retained by the fused, more massive nuclei. This implies that once a star has exhausted the fuel in its core, the thermal pressure previously generated from nuclear fusion, which prevents the gravitational collapse of the star, is no longer generated. The core becomes so hot that photons have enough energy to initiate photodisintegration, a highly endothermic process in which heavy nuclei are disintegrated into lighter elements such as hydrogen and helium. Degeneracy pressure then supports the star: the particles inside the core become more and more compressed and electrons must obey the Pauli Exclusion Principle (Pauli, 1926). The ongoing gravitational compression will supply more energy to electrons and protons, which will then combine to form neutrons while releasing a large amount of energy in the form of neutrinos. A rapid collapse of the star follows this reduction in the electron degeneracy pressure. The contraction of the core under the gravitational force continues until it is balanced by the growing neutron degeneracy pressure. The inner core further collapses until the nuclear density is so large that the strong force becomes repulsive, as a consequence of the Pauli Exclusion Principle for fermions. Subsequently, an important outgoing pressure wave is produced. Meanwhile, the outer core experiences a free fall until it encounters that pressure wave and a great flux of neutrinos produced by the continuous formation of neutrons in the inner core. This, in turns, gives rise to a supernova and leaves behind the newly born neutron star, which retains most of the rotational angular momentum of the star. If this remnant core has a mass greater than $\sim$3 $M_\odot$, the neutron degeneracy pressure will fail to support the star and it will further collapse into a black hole (Landau, 1932).

The technicalities of the core collapse supernova models are still the subject of intensive research. This is mainly because three-dimensional simulations can't quite produce a successful supernova. For example, the outer envelopes often do not fall back onto the core, therefore failing to generate an explosion. Those complex simulations are tremendously expensive computationally and they are very sensitive to the treatment of, for example, neutrino physics, energy and matter transport via convection, the role of magnetic fields, differential rotation, etc. (Couch *et al.*, 2015).

1.1.3 General properties of Neutron Stars

One of the most puzzling problems in the study of compact objects is the physical state of the matter in a neutron star. Indeed, because of the strong gravitational field of neutron stars, general relativity must be invoked when studying these objects and therefore one cannot assume that matter in the interior can be described as a typical degenerate gas (as is the case for white dwarfs). Astrophysicists have been developing models to describe the Equation of State (EoS) of the interior of neutron stars (for a recent review, see e.g., Lattimer, 2012), but those models are yet uncertain and poorly constrained. More experimental data, such as mass, surface temperature measurements and burst properties (e.g., Ransom *et al.*, 2014, Tauris *et al.*, 2011) can be useful for constraining existing models by predicting the radius that a neutron star with a certain mass should have. Based on the Chandrasekhar limit (Schönberg & Chandrasekhar, 1942), if one assumes that a typical neutron star is supported by neutron degeneracy pressure and has a radius of $\sim$10 km with a mass of $\sim$1.4 $M_\odot$, then the average stellar density would be greater than 6 $\times 10^{14}$ g cm^{-3}. This is about three times higher than nuclear density. This would imply a surface gravity ($g = \frac{GM}{R^2}$) of the order of 10^{14} cm/s^2, which is $\sim$$10^{11}$ times stronger than the gravitational force at the surface of the Earth.

Neutron stars also have the strongest magnetic fields observed in the universe: the magnetic fields at the surface typically range from 10^8 to 10^{14} G. The most accepted theory for explaining such large magnetic fields is the "flux freezing" of magnetic field lines in the conducting material inside the star. As the progenitor star collapses, the magnetic flux crossing any given open surface (not within the initial star) must be conserved. This implies that as the star shrinks, the strength of the magnetic field at the surface increases. Additional sources are sometimes required to understand the origin of the magnetic fields of the most magnetized neutron stars, called magnetars, such as more complex magnetohydrodynamic dynamo processes (e.g., Price & Rosswog, 2006).

Another important characteristic of neutron stars is their rapid rotation following the birth of the star. The formation of the compact object following the core collapse supernova is associated with an enormous reduction in the moment of inertia as the massive star collapses into a much smaller object. By conservation of angular momentum, this implies an increase in the rotational angular frequency, as shown by:

$$L = I\omega \tag{1.1}$$

where I is the moment of inertia ($I = \frac{2}{5}MR^2$ for a solid sphere of constant density) and ω is the angular frequency ($\omega = \frac{2\pi}{P}$, where P is the spin period of the star). We see from Eq. 1.1 that as the star collapses (i.e., its radius decreases), there is a reduction in the moment of inertia, which in turns causes an increase in the rotational angular frequency (since L must be kept constant). Note that the moment of inertia is dependent on the density profile, which is governed by the Equation of State of the neutron star. The variation of the predicted moment of inertia of neutron stars given existing models for the EoS are of the order of 70$\%$ (Lorimer & Kramer, 2004), but the solid sphere approximation is usually adopted for most practical calculations. Neutron stars can initially have sub-millisecond rotation periods and eventually slow down (at a spin down rate $\dot{P}$) by radiating away energy via magnetic radiation (see §1.1.4 for further details).

1.1.4 The Pulsar Model

The majority of observed neutron stars emit beams of electromagnetic radiation in the radio band. From an observer on Earth, as the neutron star rotates rapidly, this radiation will be seen as periodic pulsed emission if the beams of radio light sweep past our line of sight. The mechanism behind the radio emission is yet unknown, but we believe that it is emitted along the co-rotating magnetic axis, which can be misaligned with the rotation axis. This misalignment is necessary to observe a neutron star as being a pulsar: this is how an observer can see one pulse for every rotation of the star (or two pulses if the angle between the magnetic and rotation axis is close to 90°).

Pulsars possess co-rotating magnetospheres populated by charged particles (Goldreich & Julian, 1969) that were ripped off the surface by the strong electric field induced by the motion of that magnetosphere (Faraday's law of induced electromotive forces by varying magnetic fields). The polar cap size of a neutron star is defined by the last "closed" magnetic field line, where the latter is the field line that co-rotates with the star at the speed of light (at the stellar equator). Field lines originating from the polar cap are called "open" magnetic field lines. As electrons are accelerated and spiral along "open" field lines, they generate synchrotron radiation. This proposed theory could partially explain the non-thermal emission we see in the radio band, but that process is not yet well understood. Figure 1.1 illustrates the basic picture of the pulsar model.

As mentioned previously, pulsars lose rotational kinetic energy over time due to magnetic dipole radiation: the process in which electromagnetic waves, generated by time-varying magnetic and electric fields, carry energy away from a rotating magnetic dipole. The magnetic dipole radiation formula is as follows:

$$\frac{dE}{dt} = -\frac{32\pi^4 B^2 R^6 \sin^2\theta}{3c^3 P^4} \tag{1.2}$$

where B is the surface magnetic field, R is the radius of the star and θ is the angle between the magnetic and the rotation axis (Lorimer & Kramer, 2004). The rate at which such a rotation-powered pulsar loses kinetic energy, also called the spin-down luminosity $\dot{E}$, is given by:

$$\dot{E} = -\frac{dK}{dt} = 4\pi^2 I \frac{\dot{P}}{P^3} \tag{1.3}$$

where I is the moment of inertia, $\dot{P}$ is the spin down rate (i.e., the rate of change of the period), P is the period, and K is the kinetic energy of the star ($K = \frac{1}{2}I\omega^2$, with ω being the angular frequency). The magnetic field at the equator of the neutron star can then be estimated by equating the two expressions :

$$B = \left(\frac{3c^3 I}{8\pi^2 R^6}\right)^{1/2} \sqrt{P\dot{P}}. \tag{1.4}$$

For a canonical neutron star of mass 1.4 $\mathrm{M}_\odot$ with a radius of 10 km and where the moment of inertia is the one of a sphere with uniform density ($\sim 10^{45}$ g cm^2), the surface magnetic field is approximately:

$$B \sim 3.2 \times 10^{19} \sqrt{P\dot{P}} \text{ Gauss} \tag{1.5}$$

The magnetic field can therefore be estimated directly from two observable quantities, P and $\dot{P}$.

One can derive the expected evolution of the angular frequency (ω) from the formulas described earlier:

$$\dot{\omega} = -K\omega^n \tag{1.6}$$

where K is generally constant and n is called the braking index, a useful quantity that can give insights into the physics at play in the star. For a pure magnetic dipole rotating in vacuum, the braking index is $n = 3$. However, observational data often point toward braking indexes smaller than the canonical value $n = 3$ (e.g., Livingstone *et al.*, 2007), therefore the rotating pure magnetic dipole model is possibly an over-simplistic model of pulsars. This index can be used to determine the approximate age of the neutron star :

$$T = \frac{P}{(n-1)\dot{P}} \left(1 - \left(\frac{P_0}{P}\right)^{n-1}\right) \tag{1.7}$$

where P_0 is the spin period of the pulsar at birth, which is typically unknown. This equation also assumes that the braking index of the star has been constant since birth. If we further assume that $n = 3$ and that $P_0 \ll P$, then we obtain the characteristic age τ_c:

$$\tau_c \equiv \frac{P}{2\dot{P}}. \tag{1.8}$$

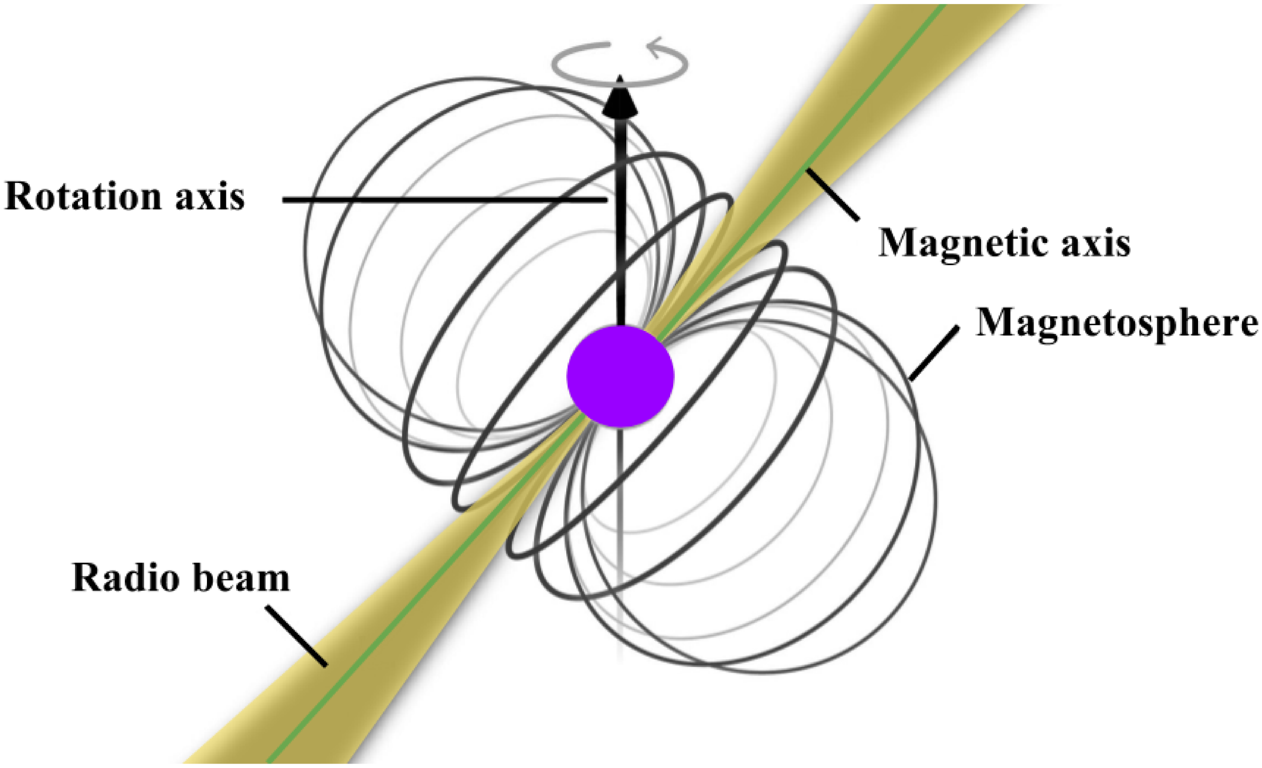


Figure 1.1: Illustration of a pulsar: a rapidly spinning neutron star. The radio beams emanate from the magnetic poles, and the magnetic axis is misaligned with the rotation axis. One can see on this figure magnetic field lines forming the magnetosphere. Image modified from Jen Christiansen (http://jenchristiansen.com/)

1.2 Observations of Pulsars

Neutron stars emit light across the full electromagnetic spectrum, and pulsations are not only observed in the radio band. They can commonly be observed as pulsars in the X-ray band, where the emission can have a thermal spectrum if the photons are emitted from the hot surface and/or have a non-thermal spectrum originating from magnetospheric processes. Some pulsars are also observed in gamma-ray as well as in optical bands, such as the Crab (Browning *et al.*, 1971, Cocke *et al.*, 1969) and the Vela (Thompson *et al.*, 1975, Wallace *et al.*, 1977) pulsars.

1.2.1 Pulse profiles in the radio band

Radio frequencies (typically between ~ 100 MHz and ~ 100 GHz) are the most frequent emission bands in which we observe pulsars, and only radio observations of pulsars will be further discussed in this thesis. The great challenge with regard to radio pulsar observations is that these stars are very weak radio sources (i.e., single pulses are generally below the noise intensity level). Most often, the coherent addition of a large number of pulses at the spin period (i.e., folding at the fundamental spin frequency) of the pulsar is required in order to detect a signal, producing what is called an integrated pulse profile. This pulse profile is composed of individual pulses differing in shapes, but the integrated profile of a pulsar is generally stable for a given observing frequency band.

Pulse shapes vary from source to source (see examples in Figure 1.2) and depend on the structure and the size of the emitted radio beam. The angle between the beam and our line of sight will also affect the observed pulse shape, as the pulse shape is dependent on both the observing frequency and the radial distance from the centre of the beam. Integrated pulse profiles at different radio frequencies for two pulsars are shown in Figure 1.3. It is believed that the majority of normal pulsars show broader pulses at lower frequencies, where the profile components tend to be separated in phase (Komesaroff, 1970). Scientists refer to this phenomenon as the "radius-to-frequency mapping" (Cordes, 1978): higher frequency emission originates from lower altitude regions closer to the neutron star surface at the magnetic poles, compared to the lower frequency components (Ruderman & Sutherland, 1975). Over larger time scales, it is possible to see an evolution in the pulse profile. This may be due to a change in the orientation of the beam with respect to our line of sight, as the pulsar undergoes geodetic precession (e.g., Kramer, 1998) or free precession for isolated pulsars (see Stairs *et al.*, 2000). Another explanation for changes in pulse profiles is that the magnetosphere of the star is transitioning into another state, referred to as the "mode-changing" (first recognized by Backer, 1970, see e.g., Wang *et al.*, 2007).

More details on pulse structure and on the properties of integrated profiles are offered in Lorimer & Kramer (2004). The next section describes the propagation effects caused by the interstellar medium (ISM) on signals from pulsars and how they affect observations done with radio telescopes.

1.2.2 Propagation in the interstellar medium

Scintillation, scattering and pulse dispersion are three propagation effects affecting a radio signal, emitted by an astrophysical source and received by an observer on the Earth. The interaction

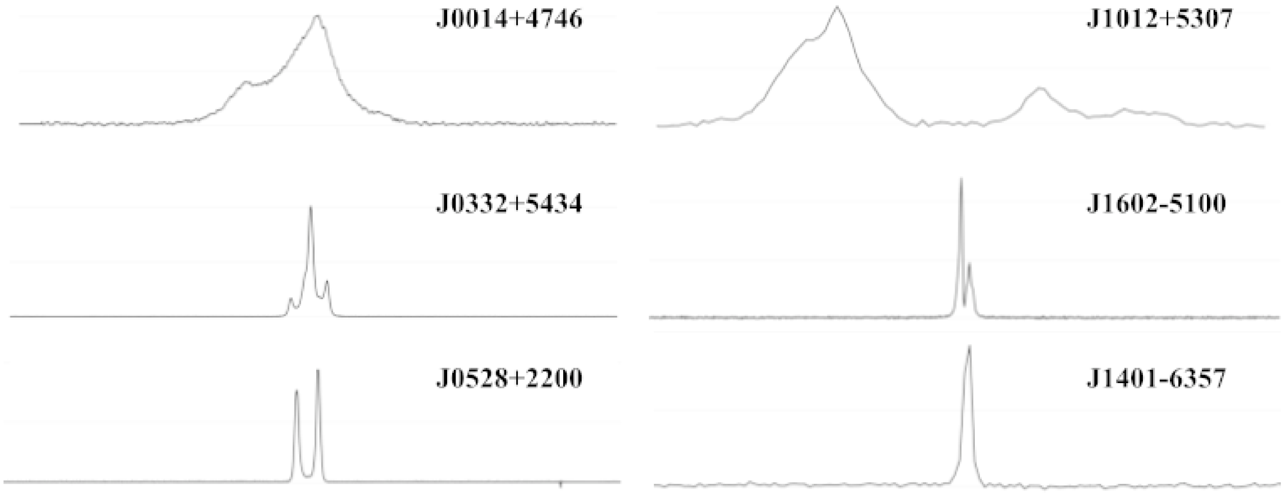


Figure 1.2: Integrated pulse profiles at 1.4 GHz for 6 different pulsars over a full phase, showing the wide variety of radio emission from pulsars. These profiles are available publicly on the EPN Database of Pulsar Profiles (http://www.jb.man.ac.uk/pulsar/Resources/epn/).

between the propagating electromagnetic wave and the dust and gas that form the interstellar medium is the source of these effects, which can make the detection of a signal challenging.

Interstellar Scintillation

The turbulent and inhomogeneous nature of the ionized ISM leads to irregularities in the refractive index in the travel path of a propagating radio wave. This gives rise to what is called interstellar scintillation (ISS), described as an apparent variation in flux density that can be seen on timescales from minutes to days. Fluctuations in the intensity, which are greater for sources with smaller apparent angular size, are correlated with a characteristic scintillation bandwidth $\Delta f \propto f^4$, where f is the observing frequency (Scheuer, 1968,Lorimer & Kramer, 2004).

Interstellar Scattering

A one-sided exponential tail in pulse profiles of distant pulsars can sometimes be observed when clumps of neutral matter in the ISM introduce variations in the light travel path length. This chromatic effect, called interstellar scattering, affects times of arrival of the pulses. Scattering affects lower frequencies more strongly: the induced time delay (referred to as the scattering time), is $\tau_s \propto f^{-x}$, where x is ~ 4 (Bhat *et al.*, 2004). Scattering time scales are typically on the order of

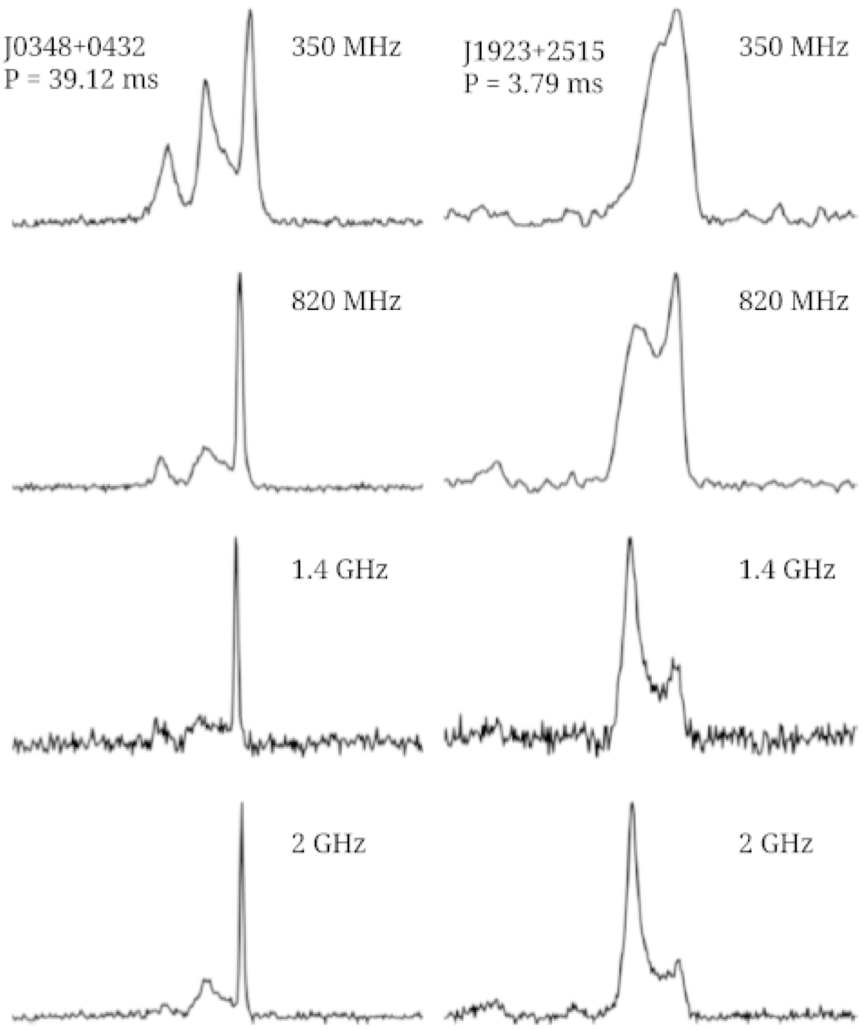


Figure 1.3: Integrated pulse profiles showing the evolution in different radio frequency bands for two different pulsars: PSR J0348+0432 and PSR J1923+2515. Figure taken from Lynch *et al.*, (2013).

milliseconds, and can be as large as a few seconds for the most heavily scattered objects, like those near the Galactic center. Unlike pulse dispersion, interstellar scattering cannot be corrected for.

Pulse Dispersion

The dispersion of the pulse is a process in which light encounters charged particles in the interstellar medium, causing a delay in the time of arrival of the lower frequencies at the telescope compared to the higher frequencies. In the presence of protons and electrons, the wave nature of the light will give rise to a delay in the frequency-dependent group velocity of the propagating signal. This is because the charged particles oscillate along with the oscillating electric and magnetic fields of the propagating light. The net result of this interaction is a delay in the phase of the propagating wave. The strength of this effect is anti-correlated with the mass of the particles and the frequency of the propagating wave. Therefore, electrons react more freely to the propagating wave compared to protons, as the latter require more energy to accelerate. Moreover, less energetic photons (i.e., lower frequency photons) are more significantly delayed compared to more energetic photons.

For a given observing frequency range, the induced time delay between the lowest and the highest frequencies is given by:

$$\Delta t = 4.1488\text{ms} \times \left(\left(\frac{f_{low}}{\text{GHz}} \right)^{-2} - \left(\frac{f_{high}}{\text{GHz}} \right)^{-2} \right) \times \left(\frac{\text{DM}}{\text{pc cm}^{-3}} \right) \tag{1.9}$$

where f_{low} is the lowest frequency of the observing band, f_{high} is the highest frequency, and the variable DM is called the dispersion measure. One can see from Figure 1.4 and Equation 1.9 that the time at which the pulse arrives follows a $1/f^2$ function. The dispersion measure, defined as the integrated column density of free electrons along the line of sight toward the source, is the quantity (in units of electrons per unit area, or equivalently in pc cm^{-3}) used to evaluate the significance of the induced time delay of the pulse. The following equation defines the dispersion measure:

$$\text{DM} = \int_0^d n_e \text{dl} \tag{1.10}$$

where d is the distance to the source and n_e is the electron number density. Cordes $\&$ Lazio (2002) have developed a map of the distribution of the free electrons in our Galaxy: the NE2001 Galactic model, where the maximum value of dispersion from our Galaxy is estimated to be $\sim$ 1800 pc cm^{-3}. However, the magnetar J1745$-$2900 in the Galactic centre has a DM of 1778 pc cm^{-3}, the largest value of dispersion measure of any known pulsar, see Eatough *et al.*, 2013). This indicates that the NE2001 Galactic model is likely underestimating the amount of plasma in the Galaxy, at least with regards to the Galactic centre.

The DM value can then be used as a distance indicator, especially useful when trying to determine if a signal is extragalactic: high DM sources tend to be located farther away or in the Galactic disk. Figure 1.5 shows the distribution of dispersion measures of the population of known radio pulsars, including those observed in the Magellanic Clouds, the closest neighbouring (dwarf) galaxies.

For a typical Galactic pulsar (DM $\sim$ 100 pc cm^{-3}) observed with the PALFA pulsar survey (central observing frequency of 1375 MHz with a bandwidth of 322 MHz, see Chapter 2), the time delay caused by the pulse dispersion by the interstellar medium would be:

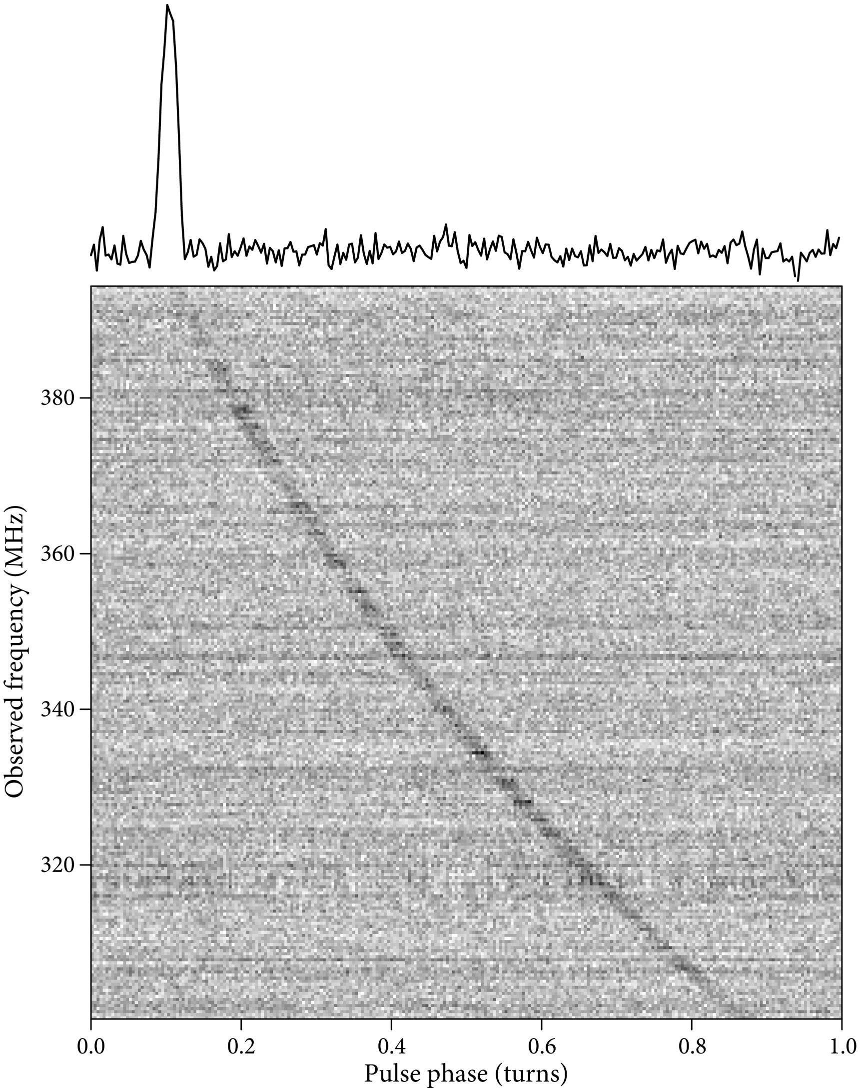


Figure 1.4: Pulse dispersion caused by the interstellar medium. *Top*: Band-integrated pulse profile once the effect of pulse dispersion has been removed (i.e., the data have been dedispersed). *Bottom*: Observed frequency delay (uncorrected), once the data have been folded at the spin period of the signal. Note the $1/f^2$ sweep of the pulse. Figure taken from Condon & Ransom, (2016).

$$\Delta t = 4.1488\text{ms} \times \left(\left(\frac{1.214}{\text{GHz}} \right)^{-2} - \left(\frac{1.536}{\text{GHz}} \right)^{-2} \right) \times \left(\frac{100\text{pc cm}^{-3}}{\text{pc cm}^{-3}} \right) \sim 106 \text{ ms.} \qquad (1.11)$$

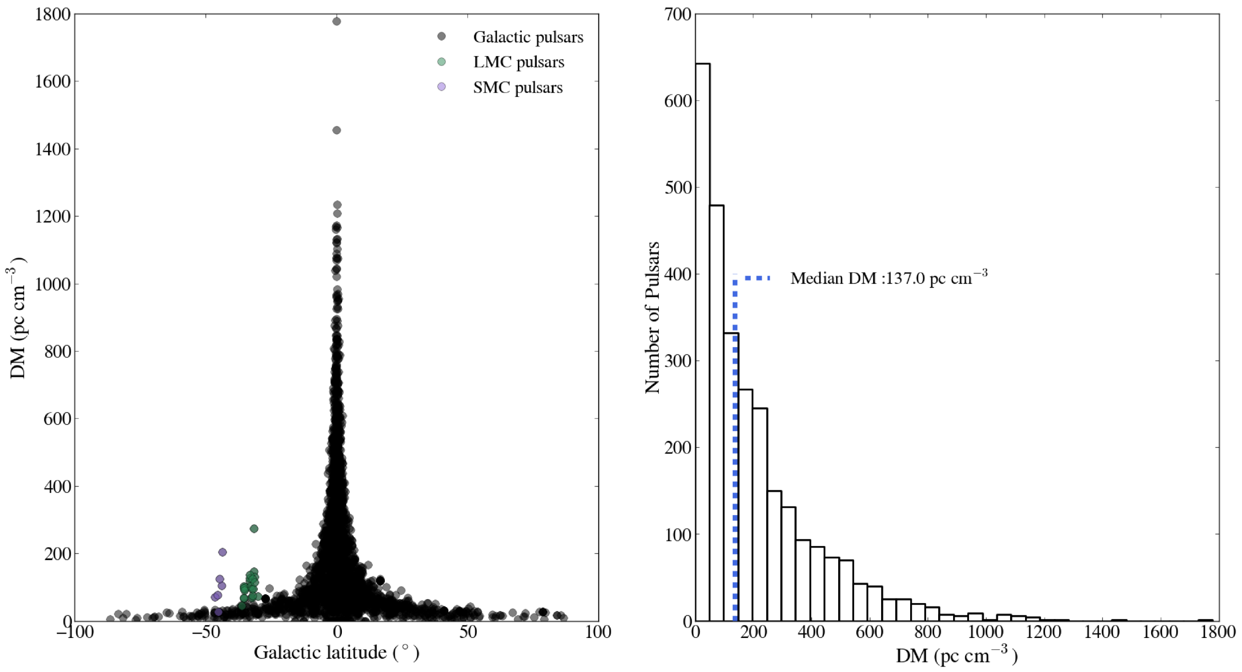


Figure 1.5: *Left panel*: Distribution of the DM values of known radio pulsars versus Galactic latitude. Green and purple circles are pulsars found in the Large and Small Magellanic Clouds, respectively. Note that the dispersion measures are larger at low Galactic latitude (Galactic latitudes near 0° on the left panel), as expected since the gas and dust densities are greater in the Galactic disk. The pulsar at a DM of $\sim$1780 pc cm^{-3} is PSR J1745-2900 (Mori *et al.*, 2013), the magnetar in the Galactic centre orbiting the Milky Way's supermassive black hole, Sagittarius A*. *Right panel*: Histogram showing the DM distribution of known radio pulsars, where the median of the distribution is identified by the blue dotted line (median value of 137 pc cm^{-3}).

To correct for the dispersion of the pulse, one bins the bandwidth and shifts the lower frequency channels to earlier times (in the case above, the lowest frequency channel would be shifted by 106 ms) such that the times of arrival of the pulse in the different channels are the same. See §2.3.2 for more detail on how we correct for pulse dispersion when we analyze radio data.

Even if the dispersion of pulses makes detection of new sources much more arduous, the characteristic $1/f^2$ signature is an advantageous indicator that an observed source is astrophysical and not man-made radio-frequency interference (RFI). Terrestrial interference, which generally peaks at DM values of 0 pc cm^{-3}, does not encounter a plasma of charged particles as it travels toward a radio telescope: therefore such signals are not dispersed.

1.2.3 Population properties

The majority of the 2723 known radio pulsars are located in the Galactic disk (as illustrated in Figures 1.5 and 1.6), and only a small number in the Magellanic clouds (e.g. Ridley *et al.*, 2013) and in globular clusters. Due to multiple selection effects, such as the small fluence of radio pulsars, this sample is heavily biased toward bright, nearby objects. The fact that the concentration of pulsars is greater in the Galactic plane is consistent with the fact that more massive stars (i.e., neutron star progenitors), are also born in regions of low Galactic latitude. Young pulsars are furthermore confined to low latitude regions in the disk, while older pulsars can be found at greater latitude in the Galactic plane. This can be explained by the fact that neutron stars are born with high velocity following the core collapse supernova: older pulsars have had time to travel away from the disk.

The number of pulsars discovered is much smaller than the 10^6 active pulsars that are expected to populate the galaxy (Lorimer & Kramer, 2004). The low brightness of the radio signal that pulsars emit, added to the effect of the interstellar medium on the propagation of the light, can partially explain why pulsar surveys have detected such a small fraction of the predicted Galactic population. Moreover, it could be possible that there are more pulsars with radio beams with small angular diameter (i.e., with narrow pulse widths) than expected, making a detection more difficult since the probability that the narrow beam of a pulsar sweeps our line of sight is lower.

A small fraction of pulsars can be found in binary systems; isolated pulsars are much more abundant. If a neutron star has a companion star prior to its birth, it is probable that the high velocity the neutron star gains as it forms from a core collapse supernova will be sufficiently large to disrupt the binary system (e.g. Radhakrishnan & Shukre, 1985). The few percent of neutron stars that remain in a non-disrupted binary may eventually accrete matter from their companion as the latter evolves in the main-sequence, depending on the orbital parameters of the system. Such an accretion process transfers angular momentum from the donor star onto the neutron star, spinning it up to a higher spin frequency (spin periods of less than 45 ms, in general). Spun-up neutron stars, called millisecond pulsars (MSPs) or recycled pulsars, form a distinct population of older pulsars: the accretion process not only shortens the spin period of the star, but also greatly reduces the strength of its magnetic field (Bisnovatyi-Kogan & Komberg, 1974). The weak magnetic fields of millisecond pulsars are a distinguishing characteristic when classifying pulsars with short spin periods as being newly born neutron stars versus recycled pulsars.

A useful graphical representation of the population of pulsars when studying the evolution of pulsars is the plot of pulsar periods (P) and their slow-down rates ($\dot{P}$), on logarithmic scales. Figure 1.7 is an example of such a plot, referred to as the P-$\dot{P}$ diagram (or B-P diagram). In this Figure, one can clearly see that millisecond pulsars, located in the lower left corner (represented as red dots), form a separate class of objects. Magnetars, which are neutron stars with incredibly large magnetic fields, are represented as green circles and typically have spin periods longer than normal radio pulsars (periods of a few seconds). On this $P - \dot{P}$ diagram, a regular radio pulsar is born in the upper left corner, and as it ages, it moves on lines of constant B-field toward the lower right corner (although it is possible that other processes cause a decay of the magnetic fields). At late times (10^8 - 10^9 years), normal pulsars will have slowed down enough that they will cross the "death line": a curve in the $P - \dot{P}$ diagram below which the pulsar enters a new state in which the radio emission of the neutron star is no longer detectable (see §3.1 for more details on the death line).

1.2.4 Pulsars as laboratories

Pulsars are incredible tools to study physical processes that are not reproducible in laboratories on Earth. The density of matter in the interior of neutron stars is the most extreme that we know of, and we know very little about the physics of such conditions. Rotational instabilities in pulsars (called glitches) can give us insights on the properties of the material in the interior of the star and the physics at play in such severe environments. The emissions from the pulsar magnetosphere are also great tools to study plasma physics at high densities in the strongest magnetic fields observed in the Universe. The processes occurring in the material surrounding neutron stars are still poorly understood and reconciling observations and theories remains challenging even after decades of intensive research with high-quality data.

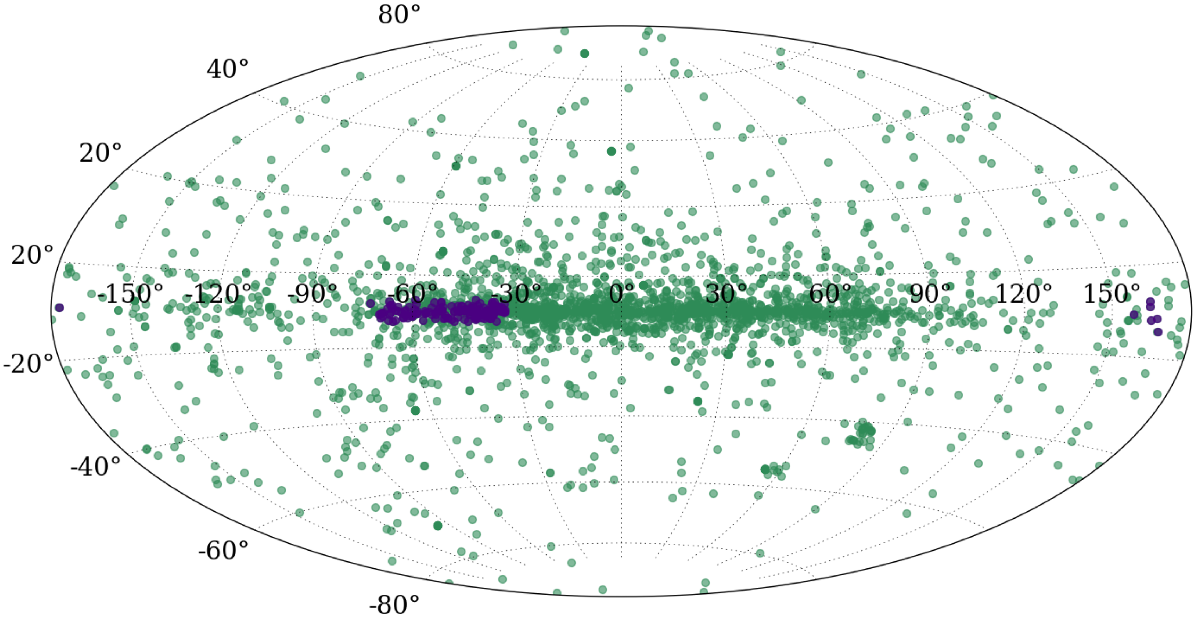


Figure 1.6: Known radio pulsar sky map in Galactic coordinates. Purple circles are objects that were discovered by the PALFA survey (see Chapter 2), while the green circles are all the other known radio pulsars.

The exquisite precision to which we can measure pulsar parameters (such as the period and the period derivative) allows us to study gravitational physics in the strong-field limit. Pulsars in binary systems, such as the Hulse–Taylor binary (Hulse & Taylor, 1975), which provided the first measurement that validated Einstein's theory of general relativity, form ideal systems to study theories of gravity and allow us to measure neutron star masses, which is essential to constrain the Equation of State of ultra dense matter. Binary systems can also reveal important pieces of information when trying to construct a consistent model of stellar evolution, particularly challenging when trying to explain how accreting pulsars can be found in eccentric orbits (e.g., Champion *et al.*, 2008, Freire *et al.*, 2008). Moreover, the exceptionally stable rotation of millisecond pulsars makes the time of arrival of individual pulses highly predictable. For this reason, MSPs can be used to form a pulsar timing array network, an alternative tool to detect gravitational waves that

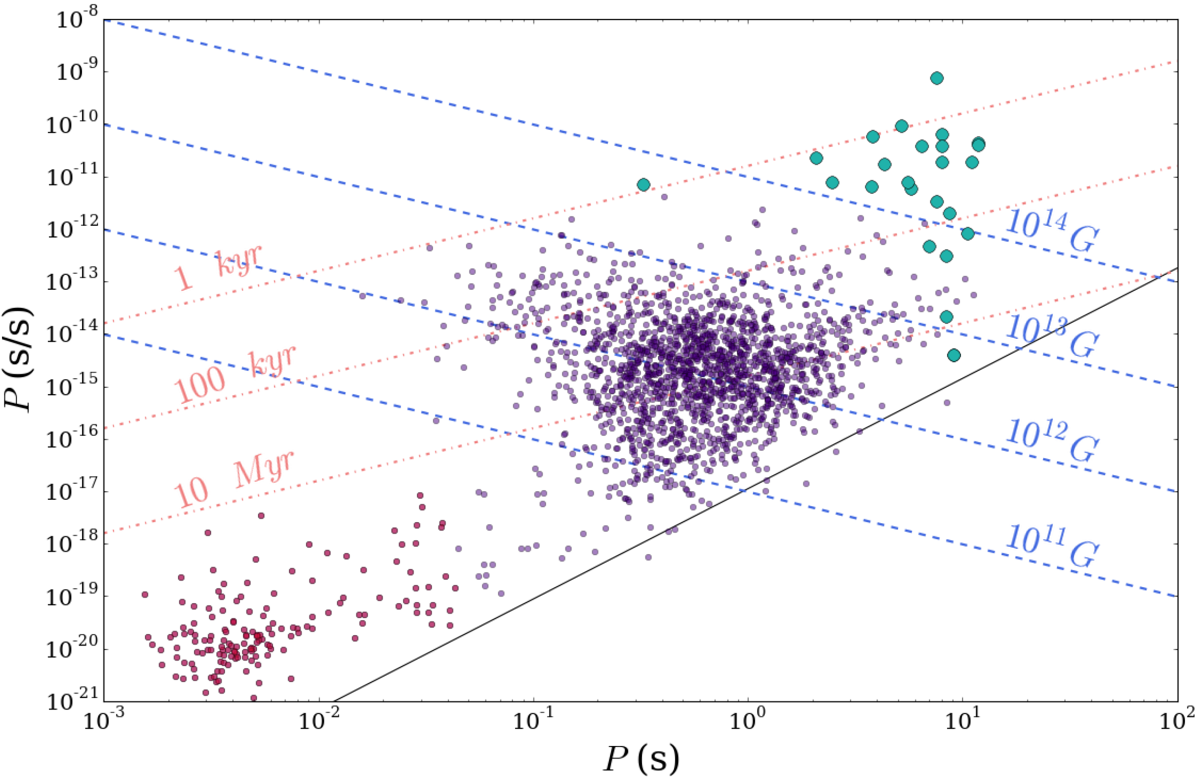


Figure 1.7: P-$\dot{P}$ diagram showing the period versus the period derivative of known pulsars (radio and X-ray pulsars). Dashed blue lines are lines of constant magnetic field (see Eq. 1.4), while the dash-dotted orange lines are lines of constant characteristic age (see Eq. 1.8). Millisecond pulsars are represented as pink circles, magnetars as green circles and normal pulsars as purple circles. The black line is one model for the pulsar death line as described in Zhang *et al.* (2000).

uses the fact that the distortion of space-time due to a passing gravitational wave will modify the path length of the photons emitted by the pulsar, hence affecting the measured times of arrival of pulses (e.g., Henning *et al.*, 2010).

Pulsars can also tell us much about our Galaxy and celestial mechanics. Indeed, to study a particular pulsar and obtain a precise timing position for that object, we need to correct for the Earth's motion in the solar system and use a reference frame centred on the Solar System barycentre. We can then use precise positions of pulsars to study, for example, physical properties of globular clusters (such as their potential gravitational well), pulsar proper motions, and their spatial distribution in the Galaxy. Dispersion measures of pulsars also help us construct maps of the Galactic distribution of free electrons. Moreover, radio signals from pulsars are linearly polarized, and their polarization will undergo Faraday rotation induced by magnetic fields as the photons travel toward us. Measuring rotations in the position angle of the signals' polarization can then tell us about the Galactic magnetic fields.

These are only a few examples of ways in which pulsars can be used as unique physical tools to study phenomena that cannot be reproduced in laboratories on Earth, highlighting the importance of pulsar astronomy. There are still many unanswered questions regarding the physics of pulsars. Discovering additional pulsars, especially atypical ones, is one of the best ways to develop a better understanding of extreme stellar remnants. To do so, the PALFA collaboration (see Chapter 2) decided to incorporate in the data analysis software a new type of periodicity search, the Fast-Folding Algorithm (FFA, see Chapter 3), while seeking new radio pulsars in the Galactic plane. This thesis describes my contribution to the implementation of this algorithm, designed to find long-period pulsars, in the PALFA analysis pipeline.

1.3 Outline of Thesis

This thesis focuses on the Fast-Folding Algorithm (Staelin, 1969), an alternative method to find periodic signals in the time-domain that folds time series at multiple trial periods in a special manner, and looks for statistically significant pulse profiles. As part of my MSc thesis, I have implemented this algorithm such that it searches for long-period pulsars in radio observations. It is now incorporated into the PALFA survey analysis pipeline, a series of programs that process radio observations made with the 305-m telescope at the Arecibo Observatory and search for both periodic and sporadic astrophysical signals.

Chapter 2 summarizes the PALFA survey and how the searches for astrophysical signals are carried out. Then, a description of the Fast-Folding Algorithm (FFA) is provided in Chapter 3, as well as the details regarding its implementation in the analysis pipeline. This chapter also provides scientific motivation for finding long-period pulsars. A sensitivity analysis of the FFA using injected signals in actual survey data is presented in Chapter 4, along with a comparison to the sensitivity of the existing FFT-based search pipeline. Finally, Chapter 5 concludes by offering a summary of the work presented in this thesis and by speculating about possible directions that future work may take.

Chapter 2

The PALFA Survey

The Pulsar Arecibo L-band Feed Array (PALFA) Consortium conducts the most sensitive, deep radio pulsar survey of low Galactic latitudes operating at 1.4 GHz. This large-scale survey started in 2004 and is expected to find hundreds of pulsars. PALFA observations are conducted with the 7-beam Arecibo L-band Feed Array (ALFA) receiver of the Arecibo Observatory William E. Gordon 305-m Telescope, in Puerto Rico.

As of February 2017, the survey has discovered 172 pulsars, including 30 millisecond pulsars (MSPs) (defined as pulsars with spin periods shorter than 45 ms) and 14 Rotating RAdio Transients (RRATs) (see Figure 2.1). This represents a $\sim 66\%$ increase in the total number of Galactic pulsars found in the surveyed region[1] (see §2.2). The unrivaled sensitivity of PALFA allows us to explore the Galactic plane to far greater distances compared to other pulsar surveys. The high frequency and time resolution, as well as the huge collecting area of the 305-m telescope, make PALFA especially powerful for finding distant (i.e., large dispersion measure) MSPs, faint pulsars, relativistic binaries as well as transient sources (RRATs and FRBs, for example).

Important results from the survey include the discoveries of PSR J1906+0746, a pulsar in a relativistic binary with the second shortest known orbital period (van Leeuwen *et al.*, 2015), the first eccentric binary millisecond pulsar, PSR J1903+0327, in the Galactic plane (Champion *et al.*, 2008), the multiple detections of rotating radio transients (Deneva *et al.*, 2009) and the recent,

---

[1] Based on the ATNF Pulsar Database (http://www.atnf.csiro.au/people/pulsar/psrcat/)

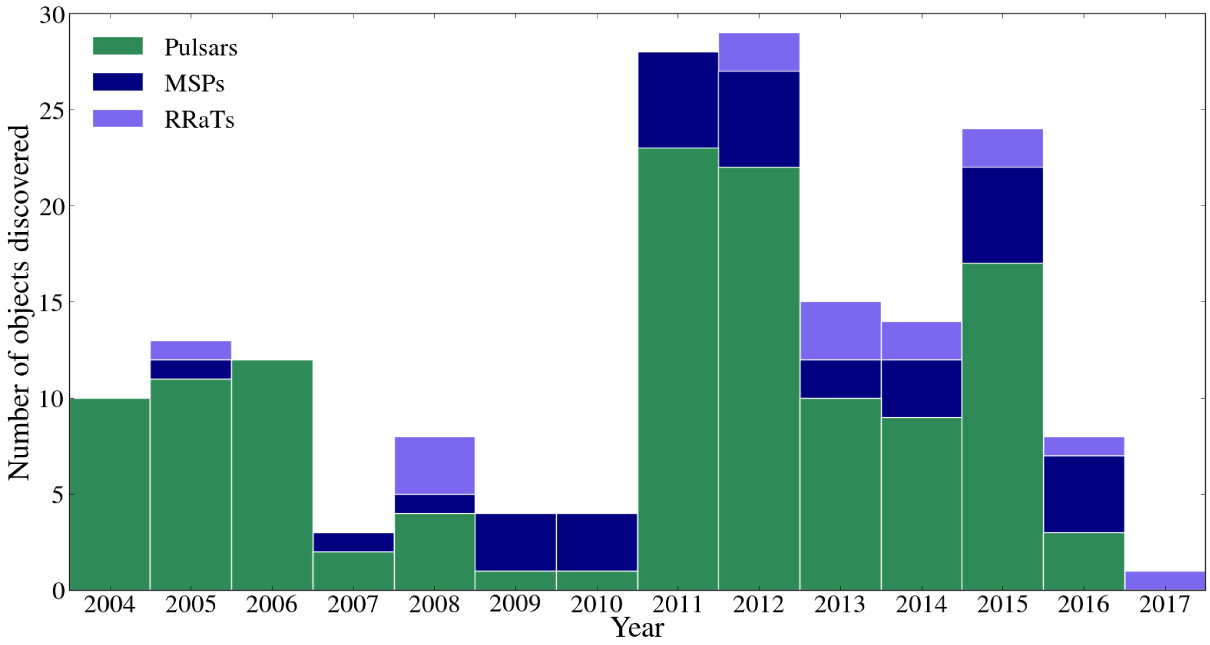


Figure 2.1: Number of PALFA discoveries per year, including normal pulsars, millisecond pulsars and rotating radio transients.

exciting discovery of the first and yet unique repeating Fast Radio Burst (Spitler *et al.*, 2016, Scholz *et al.*, 2016).

2.1 The instruments

The 305-m diameter telescope at the Arecibo Observatory in Puerto Rico, operational since 1963, is the world's largest fully operational telescope, making it the best tool to detect weak astrophysical radio signals. This telescope is equipped with a Gregorian optical system, located inside a steerable Gregorian dome suspended at the focus, 133 meters above the primary reflector (which is stationary), as seen in Figure 2.2. This primary reflector has a spherical shape with a radius of curvature of 265 m and a collecting area of 73 000 m^2. Rather than focusing the light to a focal point, as any parabolic telescope does, a spherical telescope focuses the light along a line. This necessitates a complex feed arrangement, limiting not only the telescope's observing frequency bands but also its flexibility. This is why the Gregorian optics was adopted in the late 90s: it collects the light into a single focal point while correcting for spherical aberrations and allowing observations in a much wider range of frequencies. This Gregorian optical system consists of secondary and tertiary specially shaped sub-reflectors designed for single pixel feeds, like the ALFA multibeam feeds.

ALFA, the receiver used in the PALFA survey, operates within the L-band with frequencies ranging from 1225 to 1525 MHz. It is composed of seven cooled dual linear polarization feeds in a hexagonal configuration with a half-power beam-width of roughly 3.3' $\times$ 3.7' (azimuth $\times$ zenith angle). The asymmetric side lobes, which are the regions outside the beams having non-zero gains, follow an Airy disk pattern. See Figure 2.3 for a representation of the power pattern of the ALFA multibeam receiver.

From 2004 to 2009, the WAPP Spectrometer was used to conduct the PALFA observations, providing 100 MHz of bandwidth centred at 1.4 GHz (Dowd et al., 2000). Since 2009, the survey started using the Mock spectrometers instead, which have much higher frequency resolution with a larger bandwidth of 322 MHz centred at 1375 GHz. The data are recorded via two overlapping bands of 172 MHz bandwidths (Mock 1 and Mock 2). These bands provide 960 frequency sub-channels across the full pass-band of each polarization channel (each frequency sub-channel has a bandwidth of 0.3358 MHz). This instrument also provides a high time resolution: the frequency sub-channels are sampled every 65.5 μs and the data are recorded to disk in a 16-bit search-mode PSRFITS format (Hotan *et al.*, 2004) for future off-line processing. The characteristics described above make observations with the Mock Spectrometers more robust to radio frequency interference while increasing the sensitivity of the survey.

2.2 Observations

The Galactic longitude accessible by the Arecibo telescope ranges between $32^\circ < \ell < 77^\circ$ and $168\ ^\circ < \ell < 214^\circ$, and PALFA focuses on two low latitude regions (|b| <5°) of these Galactic longitudes (see Figure 2.4). These two zones are referred as the "inner Galaxy" (toward the centre of the galaxy, at lower Galactic longitudes) and the "outer Galaxy" (toward an anti-centre region of the galaxy, at larger Galactic longitudes). The integration times of the observations are 268-sec and 180-sec for the inner and outer Galaxy regions, respectively.

Figure 2.2: Arecibo Observatory William E. Gordon 305-m Telescope. One can see the steerable Gregorian dome suspended above the primary dish, which has a secondary and a tertiary reflector that focus the light into a single point, at the location of the ALFA instrument.

To optimize the use of the telescope time, PALFA collaborates with compatible projects that use the same instrument's configuration by sharing data. Those partners are mainly looking for galaxies that are optically obscured by the Galactic plane (e.g. the Zone of Avoidance (ZOA) survey, Henning et al. 2010). Observations of the anti-centre region are led by our collaborators while PALFA is responsible for the inner Galaxy sessions.

For the calibration of the data, a noise diode is turned on for the first few seconds of the observations, and this signal is then removed prior to searching for astrophysical signals in the data. Lazarus et al., 2015 (hereafter L15) explains in more detail how this is accomplished. About 30 % of the total 450 and 460 square-degrees that are to be covered by PALFA toward the centre of the Galaxy and the anti-centre region have been surveyed. The regions with low Galactic latitudes

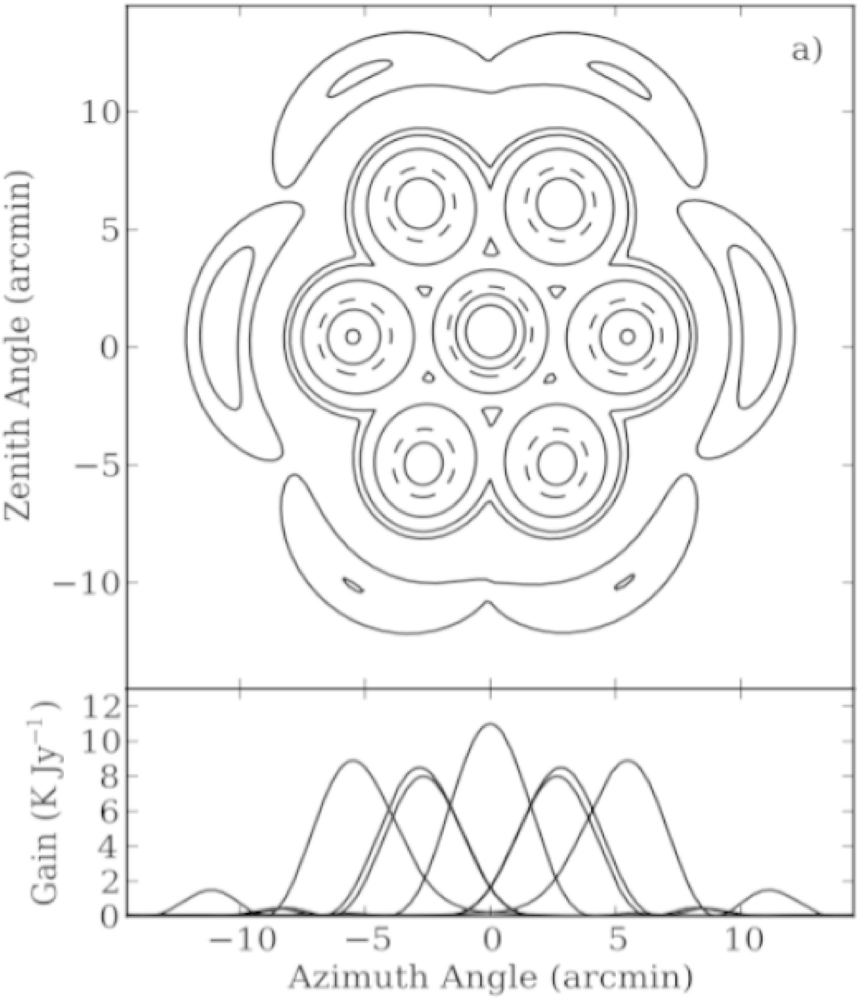


Figure 2.3: Power pattern of the ALFA receiver from the model described in Spitler et al. 2014. Top panel : the contour levels represented are -1, -2, -3, -6, -10, and -13 dB. Bottom panel: slices in azimuth for each of the 7 beams, where the slices individually pass through the peak gain of its respective beam (Figure taken from Spitler et al. 2014).

($|b| < 2^\circ$) are prioritized in the inner Galactic sector to emphasize the search in the Galactic plane, where the stellar density is greater. As mentioned previously (and illustrated in Figure 2.3), each ALFA pointing includes 7 beams with non-uniform gain. To cover extensively the entire sky without being affected by this unequal sensitivity across the beams of one pointing, the survey interleaves 3 ALFA pointings for each sky position. This ensures that the targeted region of the sky is densely sampled and makes sure that the search is uniform. Figure 2.5 shows a sky map with the

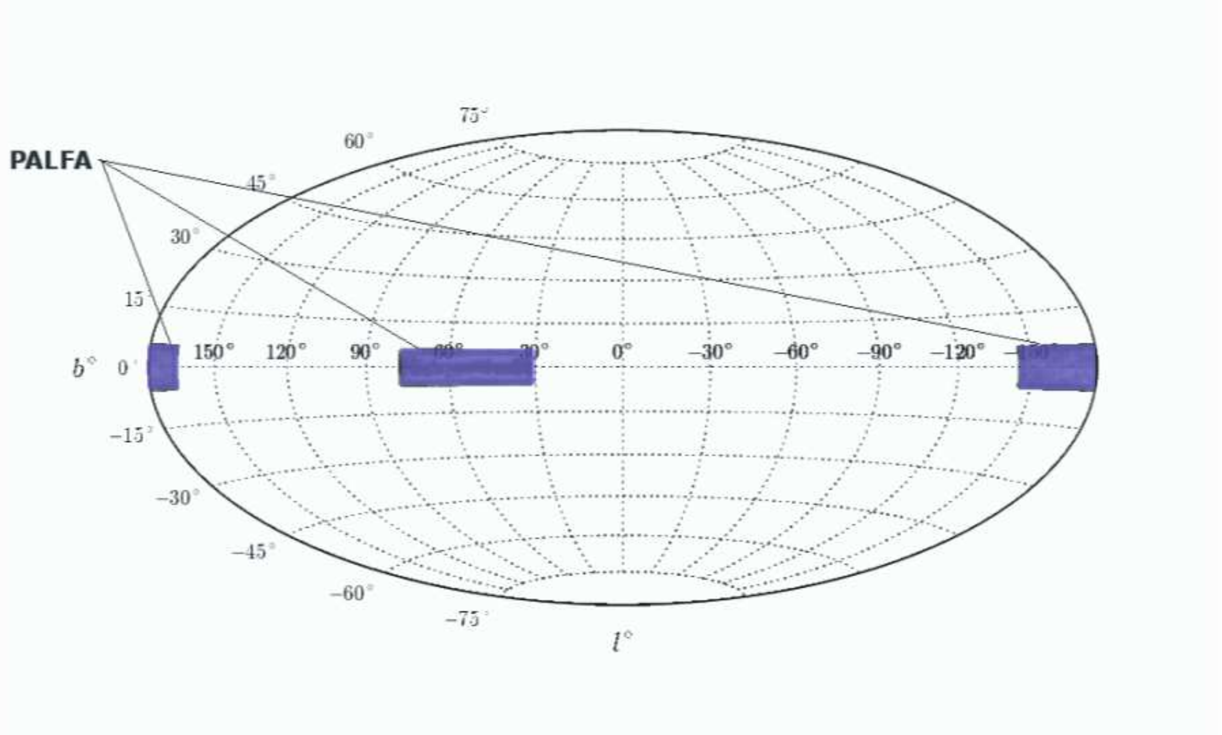


Figure 2.4: The regions of the sky covered by the PALFA survey are highlighted in violet. Image credit: http://www.mpifr-bonn.mpg.de

locations on the sky of the PALFA observations taken since 2009, when the survey started using the Mock Spectrometers. On this Figure, the positions of the centres of the sets of the 3 pointings for a same sky position are represented by the grey circles.

2.3 The PALFA Pipeline

PALFA has two independent search pipelines performing a full-resolution analysis: the primary `PRESTO`-based (Ransom 2001) pipeline and the Einstein@Home-based (Allen et al., 2013) pipeline. There is also a reduced-resolution analysis performed once an observation is completed on-site at Arecibo: the "Quicklook" pipeline, which allows rapid discovery and confirmation of bright pulsars. This thesis focuses on the `PRESTO`-based pipeline, as this is the one that now includes the Fast-Folding Algorithm pulsar search, the core subject of this thesis.

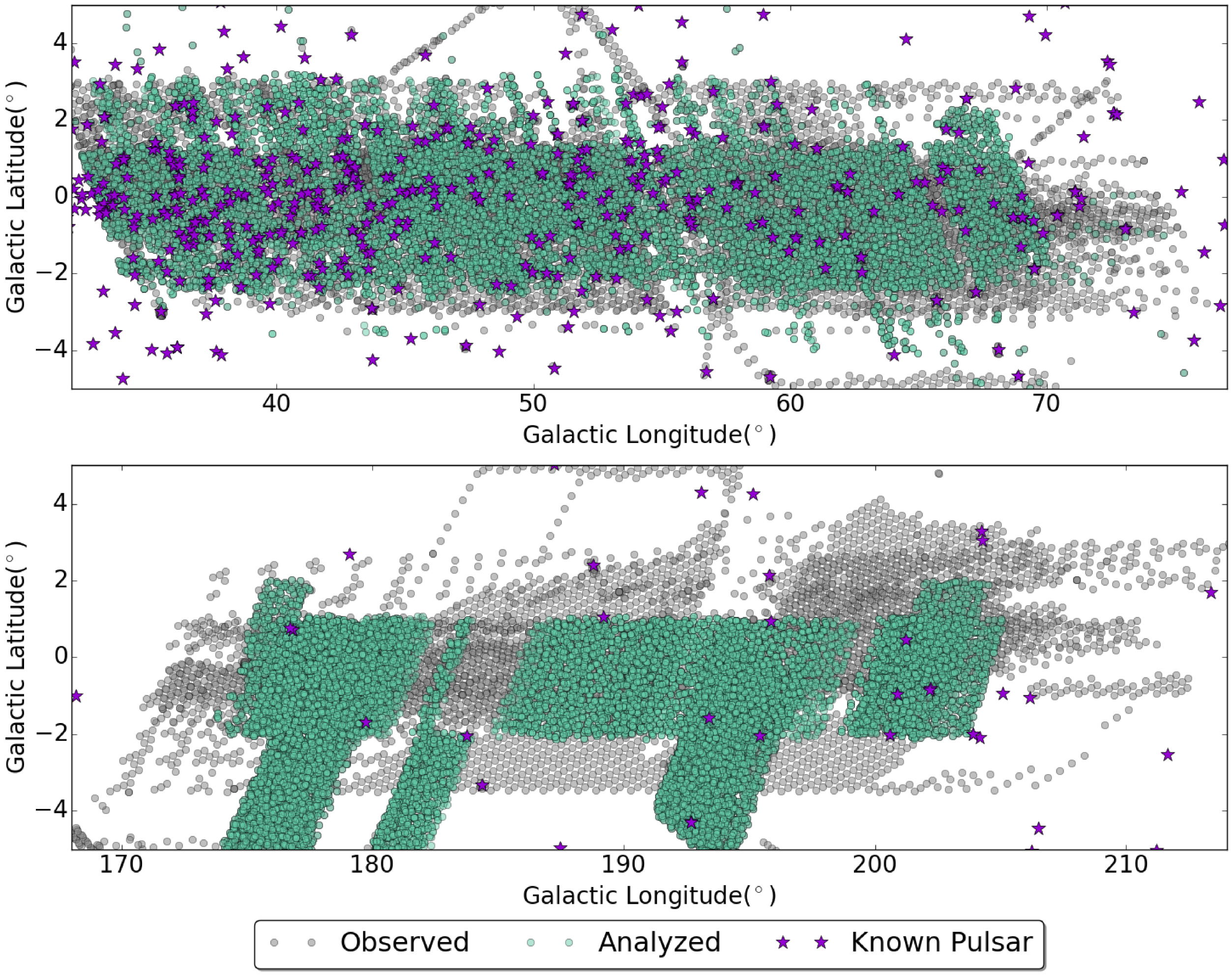


Figure 2.5: Regions of the sky covered by the PALFA survey as of February 2017. *Top panel*: inner Galaxy. *Bottom panel*: outer Galaxy. The grey circles represent the centres of the sets of the 3 interleaving pointings taken for each sky position. The green circles are the observations that have been analyzed by the PALFA pipeline, and the purple stars represent the position of known pulsars (not limited to pulsars discovered by PALFA).

The `PRESTO` pipeline runs on the Guillimin Supercomputer, part of McGill University's High Performance Computing centre operated by Compute Canada and Calcul Québec. The data are downloaded from the Cornell University centre for Advanced Computing (CAC), and the results from the search pipeline are then uploaded to the PALFA database for future inspection, also located at the CAC. Normally, the download rate is sufficiently fast to allow hundreds of jobs to run on Guillimin simultaneously. Once the processing has finished and the data has been stored at the CAC, the data are deleted from the supercomputer.

Prior to searching for astrophysical signals, one must go through a few steps to prepare the data. First of all, the two sub-bands (bandwidth of 172 MHz each) of the spectrometers (Mock 1 and Mock 2) must be combined and written to a new PSRFITS file containing 960 frequency channels with a total bandwidth of 322 MHz (some channels are overlapping). Then, the first few seconds of the observations containing the calibrating signal are removed from the file. The last stage is the generation of RFI masks to be applied to the data in order to mitigate undesirable interference (see §2.3.1). A more comprehensive description of the various steps required prior to searching for astronomical signals can be found in L15.

In this section of the thesis, the different parts of the pipeline, including the generation of these RFI masks, will be discussed.

2.3.1 Removing RFI

The PALFA pipeline includes multiple steps required to remove different kinds of ruinous terrestrial signals: some are periodic while other are spurious, and their emission can be narrow or broad band. In general, RFI signals are recognizable by the fact that they generally do not have the 1/f^2 sweep signature, characteristic of astrophysical signals (see section 1.2.1). However, some broadband interference shows frequency sweeps and are particularly difficult to identify as RFI. The interference removal routines, which are crucial parts of the analysis process, include site-specific RFI removal, Fourier-domain zapping of known interference frequencies, as well as narrow-band RFI and red noise suppression.

Site-Specific RFI Removal

The first step is to remove interference that is specific to the Arecibo Observatory, such as bursts (about 1-s long) of strong periodic signals appearing as broad features in the frequency-domain introduced by the bias monitoring system of the ALFA receiver. The collaboration has designed an algorithm capable of recognizing such signals, specific to Arecibo, and removes the affected parts of the data.

Narrow-band RFI Removal

Then, narrow-band interference is excised by examining blocks of 2-s long data for each frequency channel. The cleanliness of the blocks is evaluated based on 3 different statistics: the mean and the standard deviation of the time-domain, and the maximum value in the power spectrum. If, in one block, at least one of these three statistics is more than some standard deviations away from its respective distribution (10σ for the time-domain and 4σ for the frequency-domain), this block is flagged as contaminated. This results in a list of bad blocks that have to be masked out and replaced by the median bandpass in that frequency channel. If some channel has a fraction of masked blocks that is too large, this channel is completely replaced. The same procedure is applied for bad blocks in the time-domain: if some block has many frequency channels marked as contaminated, this block is completely deleted. If the total fraction of masked blocks exceeds 15%, we stop the processing of this observation and we later re-observe this sky position.

Time-Domain Clipping

Some broadband signals are too weak to be detected in individual channels and are consequently not masked by the process described previously. To mitigate those signals, we produce a time series dedispersed at a DM of 0 pc cm^{-3}, which should not include any astronomical sources as there is always a small quantity of free electrons between us and an astronomical source. Any time interval that statistically deviates from the surrounding time intervals is replaced by the local median.

Red Noise Suppression

Red noise also affects the power spectrum badly (see more details on red noise in §3.1). To suppress its effect from the power spectrum, we calculate blocks of mean power level for different frequency bins and then normalize the power spectrum accordingly. More details regarding the red noise suppression, and how the frequency bins are chosen can be found in L15.

Fourier-Domain Zapping

Finally, we generate a list of Fourier bins to be zapped, called the *zap list*. To create the *zap list*, we record two sets of signals: those in the DM = 0 pc cm^{-3} time series that are common to all observations made with the ALFA receiver on a given day, and those that are common to all seven pointings of an observation. The signals that appear in both these lists are added to the *zap list*, and are then excised from the data. This process is useful for reducing the damaging effects of the dynamic RFI environment (which varies on time scales from seconds to years) on the quality of our observations. In general, those zapped signals occupy about 1% to 3% of the data.

Once all those RFI-mitigation techniques have been applied to the data, which takes less than 1 % of the pipeline total processing time, we generate thousands of time series and search for astrophysical sources.

2.3.2 Dedispersion

Dispersion of an astrophysical radio signal by the interstellar medium causes a 1/f^2 sweep in a frequency-pulse phase plot, as discussed in section 1.2.1. Because the dispersion measure of sources is unknown prior to the discovery, we blindly search over 7292 DM trials. This implies that 7292 dedispersed time series are produced, with dispersion measures ranging from 0 to 10 000 pc cm^{-3}. The larger DM values are important for searching extragalactic impulsive signals (FRBs) because we expect the maximum dispersion measure from our Galaxy to be below $\sim$ 1350 pc cm^{-3}, according to the NE2001 model of the Galactic distribution of free electrons (Cordes & Lazio, 2002).

To generate the time series, frequency channels are shifted in time by a certain amount (depending on the trial DM) such that the pulses arrive at the same time for each frequency. The delayed frequency channels are then summed: the resulting vector has only 2 dimensions: time and intensity. When generating time series, one must correct for the rotation of the Earth such that the data are in the Solar System barycenter reference frame.

The `PRESTO` software (Ransom, 2001) is equipped with a program, the `DDplan.py`, that determines the best DM trials for dedispersing the raw data based on different parameters such as the central frequency, the bandwidth, the number of channels and the sampling interval. Such a plan is generated while considering different factors that contribute to the broadening of the pulses: the duration of each time sample (65.5 μs for PALFA observations), the dispersive smearing within individual frequency channels and within single sub-bands that occurs when the DM is approximated. It also includes the effect of having finite DM steps on the dispersive smearing across the entire frequency band, but it does not account for broadening due to the scattering by the interstellar medium, as this effect is impossible to correct for. From these four contributions, the total dispersive smearing at the central frequency of the total bandwidth typically ranges from sub-milliseconds (for low DMs) to tens of milliseconds for larger DMs. Approximately 5 % of the processing time of the pipeline is used for the dedispersion. For more information regarding the dedispersion plan of the PALFA pipeline, see §3.3.1 of L15.

2.3.3 Periodicity Searching

The main search component of the pipeline is a Fourier Transform-based periodicity search that is carried out on each dedispersed time series. This frequency-domain periodicity search, which is performed by `PRESTO`'s `accelsearch` program (Ransom, 2001), is divided into two main acceleration searches: a zero-acceleration search, designed to find isolated pulsars, and a search with high and low accelerations, optimized for searching for pulsars in binary systems. Isolated pulsars are recognizable by their narrow peaks in their power spectrum (at the pulsar spin frequency), while an accelerated pulsar in a binary system will exhibit broader peaks in the Fourier spectrum due to the Doppler shift induced by the relative motion of the pulsar over the duration of an observation. Ransom (2001) describes the technique used in the acceleration searches to recover the sensitivity loss due to the smearing of the peaks over Fourier bins.

Radio pulsars typically emit pulses with a duty-cycle (the pulse width divided by the period) of a few percent, meaning that the power in the Fourier spectrum is distributed between the fundamental spin frequency of the pulsar and harmonics of this frequency, where the number of harmonics can be approximated by the reciprocal of the duty-cycle. In order to make use of the power in the harmonics in standard Fourier-domain searches, we use a technique called incoherent harmonic summing: the parts of the original Fourier spectrum (e.g. the lower half) is stretched by an integer factor (e.g. 2) and then added to the original spectrum. This will directly add the amplitude of the harmonics (e.g. 2^{nd} harmonic) to the fundamental frequency, but will also increase the noise. In the PALFA pipeline, when summing the harmonics incoherently, `accelsearch` needs the stretching factors to be a power-of-two. Only 8 harmonics are summed in searches with acceleration to minimize the processing time, while the zero-acceleration search goes up to the 32^{nd} harmonic.

The significance of the candidate spin frequencies are calculated, and those with S/N larger than 2 and 3 for the zero and non-zero acceleration searches, respectively, will be recorded to a list for the post-processing stages (see §2.3.5).

The high-acceleration search is the most computationally expensive step of the Fourier Transform-based search, which takes about 42% of the total processing time, while the low-acceleration and the zero-acceleration parts occupy about 5% and 3% of the processing time, respectively.

2.3.4 Single Pulse Searching

The single pulse search, described in Patel (2016), is designed to find impulsive signals, such as rotating radio transients (RRATs) and fast radio bursts, with a matched-filtering technique (Cordes & McLaughlin, 2003) using multiple boxcar templates. This program searches for signals with pulse widths from 64 μs (the minimum sample size) up to about 100 ms. All events having signal-to-noise ratios greater than 5 will be recorded and further grouped, based on the time of arrival of the pulses and their dispersion measures (a physical pulse will likely be detected at multiple DM values and the optimal detection occurs at the true DM of the source). The list of potential single pulse candidates is then produced and will be used in the post-processing stages (see §2.3.5). The repeating fast radio burst, FRB 121102, was identified via this single pulse algorithm (Splitler et al. 2016). About 6% of the pipeline processing time is used by the single pulse search.

2.3.5 Post-Processing

Once the lists of potential candidates are in hand and the searches for astrophysical signals are completed, the multiple thousands of candidates (consisting of pairs of period-DM for periodicity candidates, or pulse time of arrival-DM for single pulse candidates) are sifted using `PRESTO`'s `sifting` (Ransom, 2001) code such that less significant, duplicate and/or harmonically related (for the periodicity search) candidates are deleted. Candidates with suspicious DM behaviour (e.g., peaking below 2 pc cm^{-3} or not appearing at consecutive DM trials) are removed, as well as candidates with periods shorter than 0.5 ms (which contribute to a large number of false-positives). More details regarding the sifting procedure in the pipeline can be found in §3.3.4 of L15.

After the sifting, there are about 300 periodicity candidates (from the Fourier-domain search only) and 250 single pulse candidates. Only the periodicity candidates having a S/N greater than 6 are folded, which is in general $\sim$60% of the candidates, and a plot similar to Figure 2.6 is produced for each of them. One can clearly see the pulse in both the time and frequency spaces, and a non-flat curve in the χ^2 versus DM subplot (i.e., this curve rises and falls around what must be the DM of the source: the DM value that leads to the optimal detection of the pulsar). `PRESTO`'s `prepfold` code (Ransom, 2001) produces these folded plots and it can execute a restrained search in the period, period derivative and DM versus χ^2 spaces to extract the optimal parameters. Regarding the single pulse candidates, a plot showing the signal in both frequency and DM space similar to Figure 2.7 will be produced for each group where at least one event has a S/N greater than 6. A single pulse diagnostic plot is generated for about 55% of the 250 candidates produced.

For both categories of plots, a number of ratings based on, for example, the shape of the profiles, the spectral information and the likelihood of being RFI, are calculated for each plot. The value of each rating is recorded into our database, along with the candidate's other parameters, such as the period, the S/N and the DM. The pipeline is also equipped with a machine-learning algorithm candidate selection, which is an image-based classification system that has the ability to recognize pulsars based on some features of their periodicity plot, generated by the `prepfold` code. The single pulse program also has a rating system, which is described in Patel (2016). All this information is helpful for subsequent human candidate classification.

Once the processing of one beam is fully terminated, the plots and the ratings are uploaded to our database located at Cornell University CAC. The last step to find astrophysical sources is the human classification of the candidates produced by the pipeline. To accomplish this task, the PALFA Consortium uses the online collaborative tool on the CyberSKA platform (Kiddle *et al.*, 2011). This infrastructure was designed especially for managing large-scale radio astronomical data in the lead-up to the Square Kilometre Array. PALFA developed applications on CyberSKA to access the database hosted at Cornell in order to view and filter candidates for the classification.

As of Fall 2016, the pipeline has been upgraded: it now includes a more refined single pulse

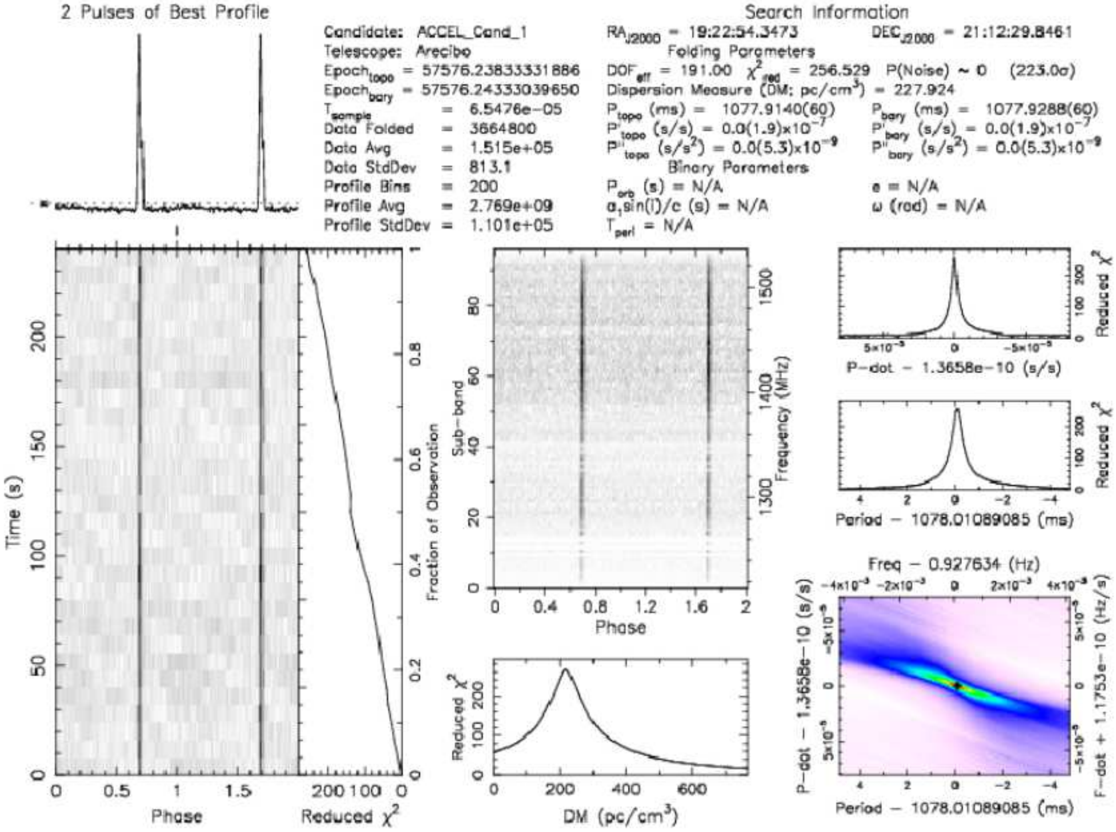


Figure 2.6: Periodicity plot generated by `PRESTO`'s `prepfold` command (Ransom, 2001) on an observation of the bright pulsar B1920+21 (spin period of 1.077 sec). General information about the pulsar is displayed at the top. *Bottom left*: intensity as a function of observing time and phase. *Top left*: pulse profile once it has been summed. *Top middle*: intensity as a function of observing frequency and phase. *Bottom middle*: reduced χ^2 versus DM. Note that the optimal DM peaks around 200 pc cm^{-3}. *Lower right*: χ^2 of a search around the period and period derivative.

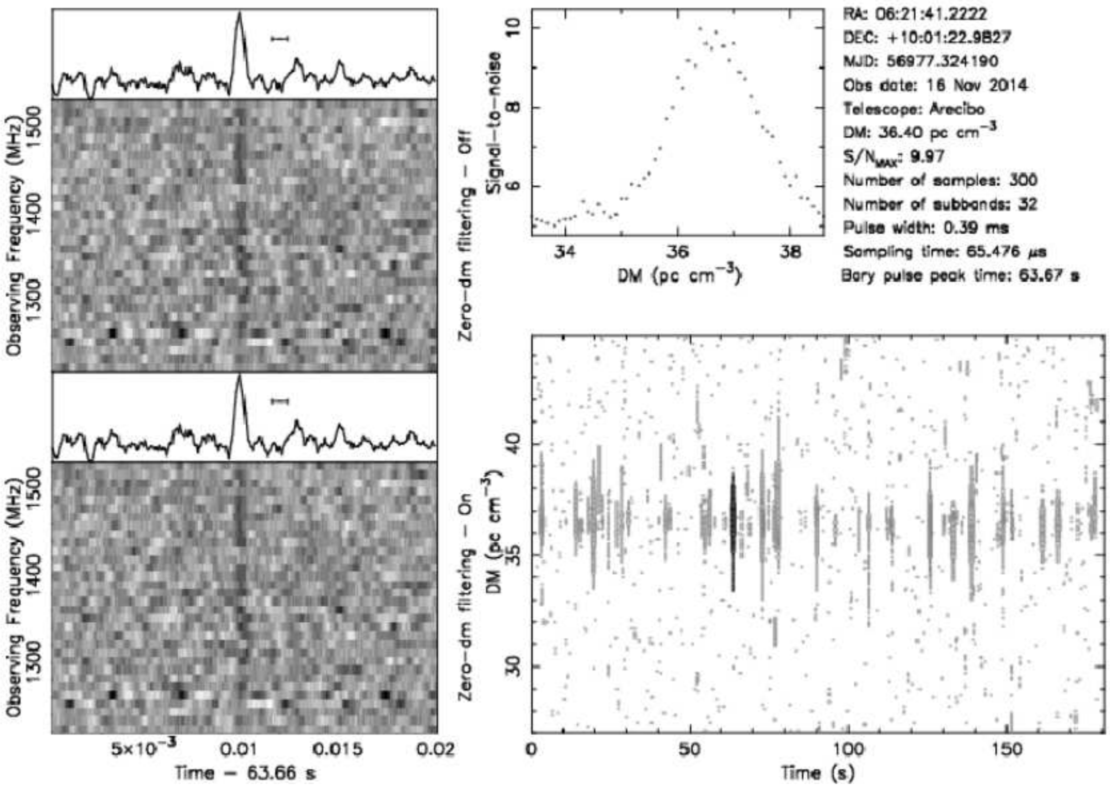


Figure 2.7: Diagnostic single pulse plot of a known millisecond pulsar J10621+1002 in a 8.3-day binary orbit with a spin period of 28.85 ms. *Left plots*: dedispersed frequency versus time of one of the pulses with (bottom) and without (top) an RFI filter (called the zero-dm filter). *Bottom right*: DM versus the observing time for all events found by the single pulse program, where the size of the points scales with the S/N of the event. The darker points on this subplot are the grouped events represented in the plots on the left. Note that many pulses were detected at the same DM (at different times) by the `single_pulse` code (Patel, 2016). *Top right*: S/N versus trial DMs with information regarding the candidate (on the right).

search (which should be more sensitive to narrow pulses at large DMs, such as FRBs), and the Fourier-based periodicity search now does harmonics summing up to the 32$^{\text{nd}}$ harmonic, rather than up to the 16$^{\text{th}}$. This modification should increase the sensitivity of the pipeline to pulsars with narrow pulses.

A major new addition to the pipeline's periodicity search is the implementation of a Fast-Folding Algorithm, which also searches for periodic signals in the time series, but looks for signals in the time-domain rather than in the frequency-domain. This implementation is the core of the research I have been doing for my MSc thesis, and Chapter 3 explains the details of this algorithm and its implementation in the PALFA pipeline. Chapter 4 shows how the sensitivity of the survey is

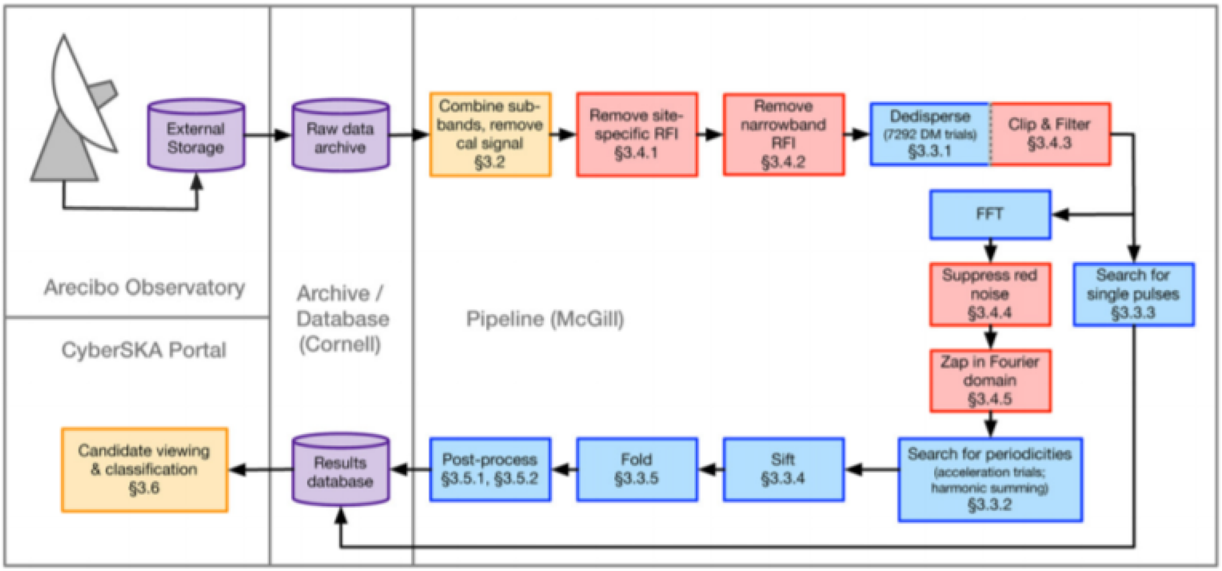


Figure 2.8: Diagram summarizing the PALFA pipeline. Purple boxes represent data storage and databases, red boxes represent parts of the pipeline that remove RFI from the data, blue boxes represent searching elements, and the miscellaneous are yellow. Figure taken from L15.

expected to increase with this new addition to the pipeline. Figure 2.8 summarizes the diverse steps involved in the pipeline.

Chapter 3

Searching for Slow Pulsars with the Fast-Folding Algorithm

3.1 Motivations

One characteristic of the population of known radio pulsars is that the vast majority of pulsars have periods shorter than 2 seconds: very few known pulsars have periods greater than a few seconds. Indeed, as shown in Figure 3.1, we can see that the median spin period of the population of known radio pulsars is $\sim$501 ms (when millisecond pulsars are included; $\sim$610 ms when excluded). The lack of long-period pulsars could be an intrinsic property of the population, but may also be due to selection bias in pulsar surveys. One of the reasons why PALFA as well as other pulsar surveys are likely to miss slowly rotating pulsars is that pulsar search data are often badly affected by red noise. This type of noise is not Gaussian distributed and has a spectral density that is approximately proportional to $1/f^2$, meaning that it is strongest at low modulation frequencies. This noise is the result of the combined effects of different factors such as receiver gain fluctuations and man-made interference. The broad features introduced in the time series by red noise increase the number of false alarms in the low spin frequency regime and increase the scatter and bias of the signal to noise ratio, meaning that our sensitivity at this end of the spectrum is considerably reduced. For multiple reasons discussed below, it is important to recover this loss in sensitivity as it has the potential to offer great scientific advancements.

First, our understanding of the Galactic pulsar population is heavily biased by various selection effects, preventing us from drawing accurate conclusions. Such biases include the non-uniform radio sky background (which dominates along the Galactic plane at observing frequencies < 0.4 GHz), the system noise temperature (which depends on the telescope), propagation effects in the interstellar medium (such as scintillation, scattering and dispersion of the signals), the distance and the proper motion of the pulsar, as well as the size of the emission beam. These effects bias radio surveys towards the brighter, closer pulsars and produce a non-uniform sample in terms of spatial distribution. In addition to the observation biases mentioned above, red noise also affects the observed population of long-period pulsars. Indeed, from Lazarus *et al.*, (2015), the PALFA survey is much less effective in finding slowly rotating pulsars because of the effect of red noise compared to what is predicted by the radiometer equation (see chapter 4). Recovering the true radio pulsar population is important for fully understanding their properties such as their evolution, their birth-rate, their luminosity and spatial distribution, and also for our understanding of classes of peculiar objects such as nulling, mode-changing, rotating radio transients and intermittent pulsars. Finding new and previously missed pulsars is one step forward in obtaining a more complete picture of the Galactic radio pulsar population.

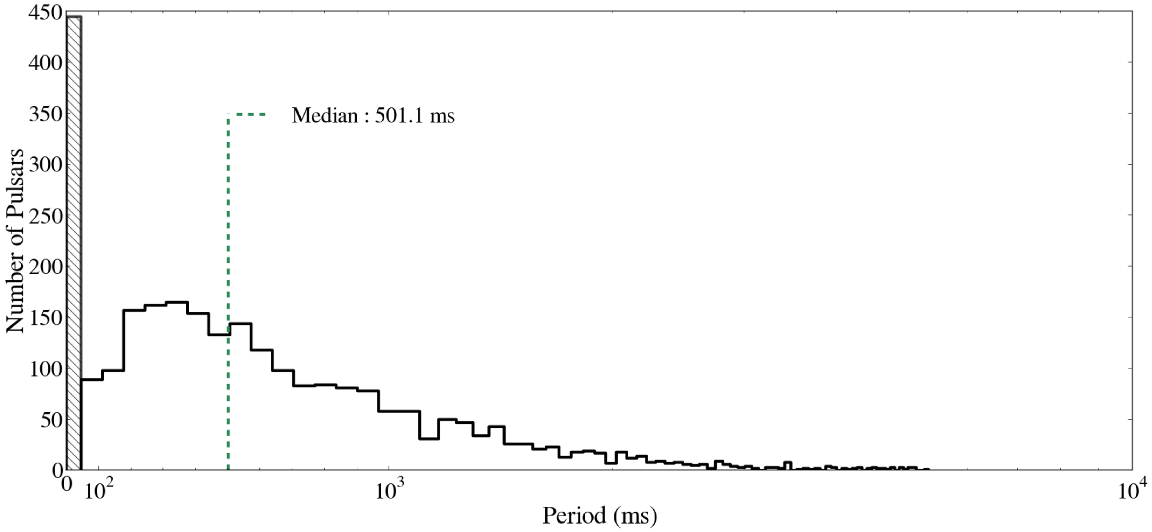


Figure 3.1: Period distribution of the 2783 known radio pulsars[1]. The dashed green line is the median of the periods (501.1 ms). Note the population of MSPs in hashed on the very left of the plot.

Second, there is yet no self–consistent theory that can explain all of the pulsars' radiation properties that we observe. Designing a rigorous model for the nature of the radio emission is challenging as it must explain the coherency of the emitted photons, the polarization, the time variability behaviours (such as drifting sub-pulses, Drake & Craft 1968; Sutton et al. 1970) and the difference in the emission spectrum from one object to another. Another complication in developing a consistent model for the radio emissions is the observed "death line", a curve in the surface magnetic field/ period (or period/ period-derivative) plane (Chen & Ruderman 1993) that separates the pulsars emitting radio light from the radio quiet ones. A transition from radio loud to radio quiet is thought to occur as a neutron star spins down with time. Eventually the electron-positron pair production, considered as an essential condition for the radio emission as it provides the required potential in order to accelerate the particles, can no longer occur in the inner magnetosphere. Given a surface magnetic field, once the neutron star has slowed down to a certain spin-period, the potential will drop and pair production will stop. This widely accepted assumption is, however challenged by the long-period (8.5 s) pulsar PSR J2144−3933, as this object is located beyond the theoretical death line, and therefore should not emit in the radio band. This implies that either the assumptions made in the standard model are wrong, or the radio emission theories are either inaccurate or incomplete (Young *et al.*, 1999). Finding more slowly rotating pulsars would help us to shine light on whether objects like PSR J2144−3933 are common, and finding slower ones would constrain even more emission mechanism models.

Third, better sensitivity to slowly rotating pulsars also increases the chances of discovering neutron star - black hole systems as we expect the spin period of a neutron star in orbit about a black hole to be long. This is explained by the fact that no accretion from a black hole companion is possible in such binary systems. Accretion is required to spin up the rotation of the neutron star as its magnetic field allows a transfer of angular momentum from the accretion disk to the neutron star. Indeed, black hole progenitors are more massive than the progenitors of neutron stars. As more massive stars evolve off the main-sequence more rapidly than lower mass stars; a black hole in a binary with a neutron star would then form first. Therefore, optimizing our detection capabilities at the long-period part of the spectrum increases the chances of discovering the first neutron star - black hole binary system, which would provide valuable insights into stellar evolution and a testbed for theories of gravity.

Finally, increased sensitivity to low modulation frequencies also makes pulsar surveys more likely to find radio-loud magnetars. Rather than emitting radiation powered by loss of rotational kinetic energy as typical pulsars do, it is thought that their extreme internal magnetic field powers the great luminosities of magnetars. Indeed, these mysterious neutron stars have the strongest magnetic fields observed in the universe ($\sim 10^{14}$ G) and they are known to have long periods. For the 25 magnetars[2] for which we know the period (Olausen & Kaspi, 2014), the median period is 6.97 sec. Out of this sample, only 4 are radio-loud magnetars, and the very nature and evolution of these objects remain unknown. Studying more radio-loud magnetars can help us understand these extreme neutron stars and constrain existing models. Moreover, it could allow us to better understand the connection between highly magnetized radio pulsars and magnetars, as magnetar-like activity from highly magnetized radio pulsars can be expected (Kaspi & McLaughlin 2005). Indeed, we know radio pulsars that show magnetar-like outbursts, such as the young pulsar PSR J1119–6127 (Archibald *et al.*, 2016) and PSR J1845−0258 (Gavriil *et al.*, 2008).

3.2 The Fast-Folding Algorithm

The Fast-Folding Algorithm (FFA) was originally developed by Staelin (1969) for searching periodic signals in the presence of noise in the time-domain, unlike the Fast Fourier Transform search technique which operates in the frequency-domain. By avoiding redundant summations, the FFA is much faster than standard folding at all possible trial periods: it performs summations through $N\log_2(\frac{N}{P}-1)$ steps rather than $N(\frac{N}{P}-1)$ where N and P are the number of samples in the time series and the folded period in units of samples, respectively. Large computational power is still required when applying the FFA over a very wide range of trial periods and this is why the FFA has never been implemented on large pulsar searches before. The use of this algorithm has been fairly limited over the last decades. Lovelace & Sutton (1969) implemented the algorithm when working at the Arecibo Ionospheric Observatory and helped find the second pulsar (spin period of 0.5579 s) in the Vulpecula constellation (Craft *et al.*, 1968). More recently, an FFA has also been

---

[2] See the magnetar catalogue http://www.physics.mcgill.ca/ pulsar/magnetar/main.html

applied in an RRAT study (Losovsky & Dumsky, 2014) which showed that periodic weak pulses were emitted from these pulsars. Crawford *et al.*, (2009) also used the Fast-Folding Algorithm while conducting a periodicity search on radio observations of the 6.85 sec X-ray pulsar XTE J0103−728, but flat search resulted in no significant candidates. Having a greater resolution in period compared to a Fourier search (as discussed further in following sections), it has been used on observations of sources where the approximate period was known, in order to have a more precise period measurement (Kondratiev *et al.*, 2009). This thesis represents the first time that the FFA algorithm is used to conduct a blind periodicity search in a large-scale pulsar survey.

The FFA folds each time series dedispersed at some specific dispersion measure with some sampling interval Δt at multiple periods and then looks for statistically significant pulses in the generated profiles, as illustrated in Figure 3.2. This algorithm performs partial summations, avoiding redundancy, into a series of $\log_2 P$ stages and then combines those sums in different ways so that it folds the data at the different periods P_i. If a time series contains N bins and the folding period in units of samples is P, where the period in time units is $p = P \times \Delta t$, the FFA will produce $M = \frac{N}{P}$ different profiles with slightly different periods in the range from P_i to $P_i + 1$:

$$P_i = P_0 + \left(\frac{i}{M-1}\right) \tag{3.1}$$

where P_0 is the effective folding period and $0 < i < M - 1$.

Figure 3.3 shows the periodogram one obtains from applying the Fast-Folding Algorithm on a 268-sec PALFA observation of the bright, long-period pulsar PSR J1901+0413 when looking for periods between 100 ms to 30 sec. The peak in signal-to-noise ratio is at the pulsar's fundamental period, 2.66 sec, and the secondary peaks are the harmonics and sub-harmonics of the period. The FFA requires $\log_2(\frac{N}{P})$ to be an integer, or equivalently, M to be a power of 2. If this condition is not satisfied, the time series will be padded by the median value of the data in the time series until this condition is satisfied.

The main advantages of the Fast-Folding Algorithm over the Fast Fourier Transform (FFT) are that the FFA offers greater frequency resolution (especially important in the low-frequency end of the

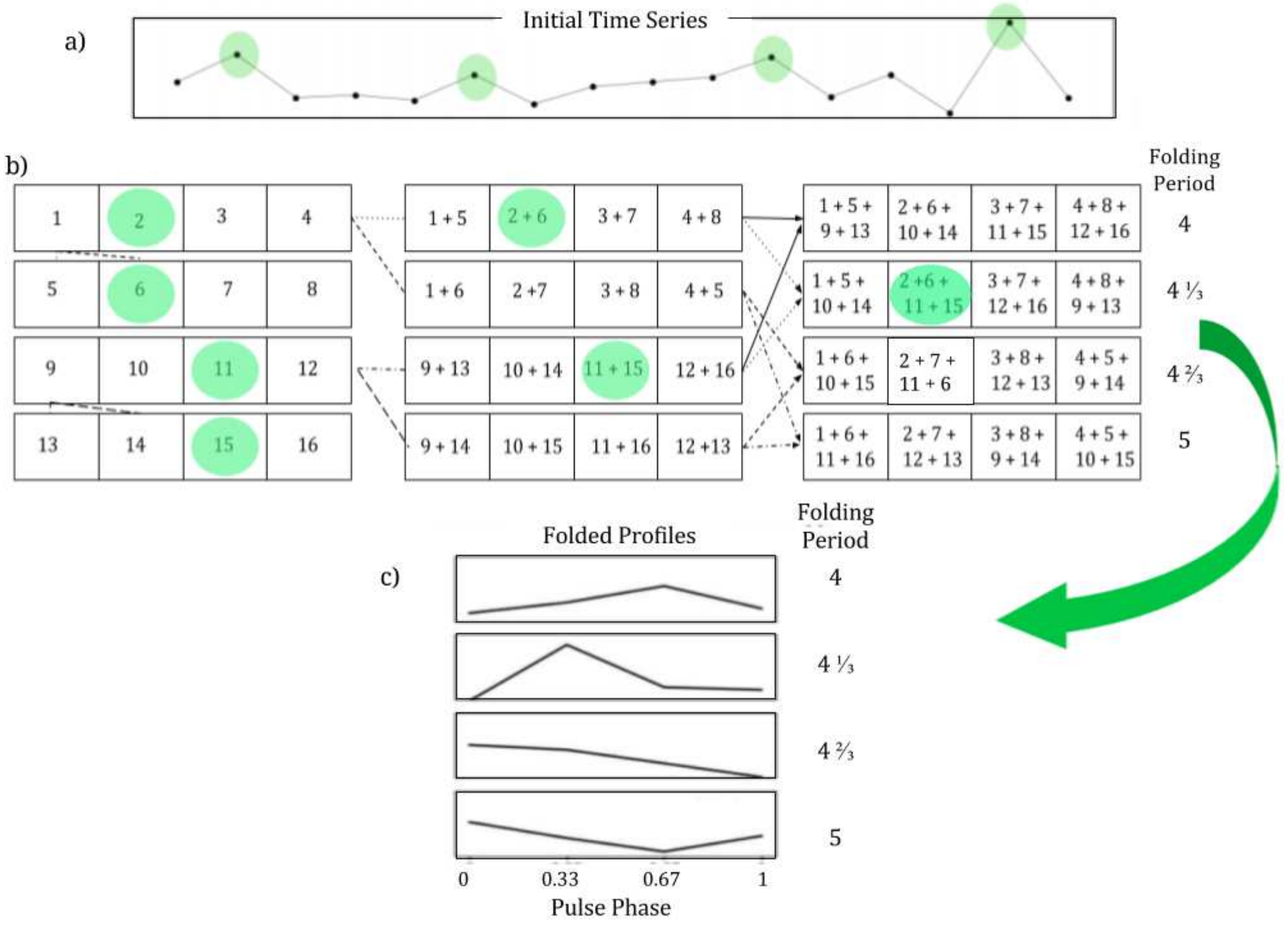


Figure 3.2: a) Initial time series containing 16 bins. The green circles illustrate the bins where the periodic signal is. b) Schematics of the Fast-Folding Algorithm applied on this time series, with a trial period of 4 bins. The summations are performed in 2 steps ($\log_2 4$) and the output is 4 (M = 16/4) folded profiles with periods: 4, 4.33, 4.66 and 5 sec. c) Resulting profiles. One can see the pulse emerging from the 2$^{\text{nd}}$ profile, meaning that the period of the signal is close to 4.33 sec.

spectrum) and, most importantly, that it sums coherently all harmonics of a signal, i.e., it folds the data in phase. Indeed, the incoherent harmonic summing that is used in Fourier-domain searches inevitably loses power in higher harmonics, as one has to choose a finite number of harmonics to be summed when using this search technique (maximum of 32 summed harmonics in the case of the PALFA PRESTO-based pipeline). All the power contained in the harmonic structure is used in the FFA, making this technique much more sensitive to narrow pulses.

The known population of long-period radio pulsars[3] (with periods greater than 2 sec) tend to

---

[3] Based on the ATNF Pulsar Database (http://www.atnf.csiro.au/people/pulsar/psrcat/)

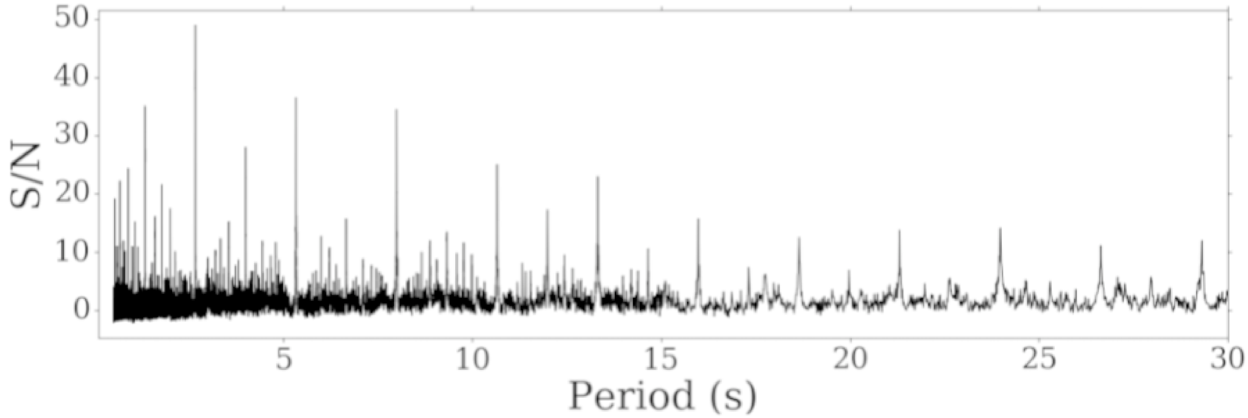


Figure 3.3: Periodogram of PSR J1901+0413. One can clearly identify the fundamental period of the pulsar, 2.663 sec, as well as multiple significant harmonics (at shorter periods than the fundamental period) and sub-harmonics (at longer periods than the fundamental period). The shortest tested period is 100 ms, and the search extends up to periods of 30 sec.

have narrow pulses, as shown in Figure 3.4. Indeed, the median pulse duty-cycle for this class of pulsars is 1.6 %, while it is 3.1 % for pulsars with spin periods shorter than 2 sec. PSR J0627+16 has a period of 2.180 sec and is incredibly narrow peaked: it has a FWHM of only 0.3 ms (Deneva *et al.*, 2009), which corresponds to a pulse duty-cycle of 0.01 %. Finding more extreme objects like PSR J0627+16, which can tell us about the mechanism behind the radio band emission of pulsars, requires a maximal use of the harmonic content. Furthermore, the longest-period pulsar known to date, PSR J2144-3933 (Young *et al.*, 1999), was initially identified (mistakenly) as a pulsar with a spin period 3 times shorter than the true period of 8.51 sec (i.e., it was detected via its 3^{rd} harmonic). Having a search technique that is efficient at finding narrow pulsed signals in the long-period regime is therefore promising and worth employing.

In Kondratiev *et al.*, (2009), the efficacy of the FFA is compared against an FFT search with a simulation including fake pulsar signals, with different periods, brightnesses and pulse widths, injected into fake time series with Gaussian noise. In this analysis, it was shown that the signal-to-noise ratio one obtains from the FFA starts being greater that the one from the FFT at a period of $\sim$2 sec for a pulse width of 1%, and at $\sim$6 sec for a 4 % pulse width. Moreover, it was demonstrated in this study that for very long periods (P $>$ 10 sec), some trials detected by the FFA search were not detected in the FFT search when summing up to the 32th harmonic, highlighting the importance of a time-domain search when looking for long-period pulsars.

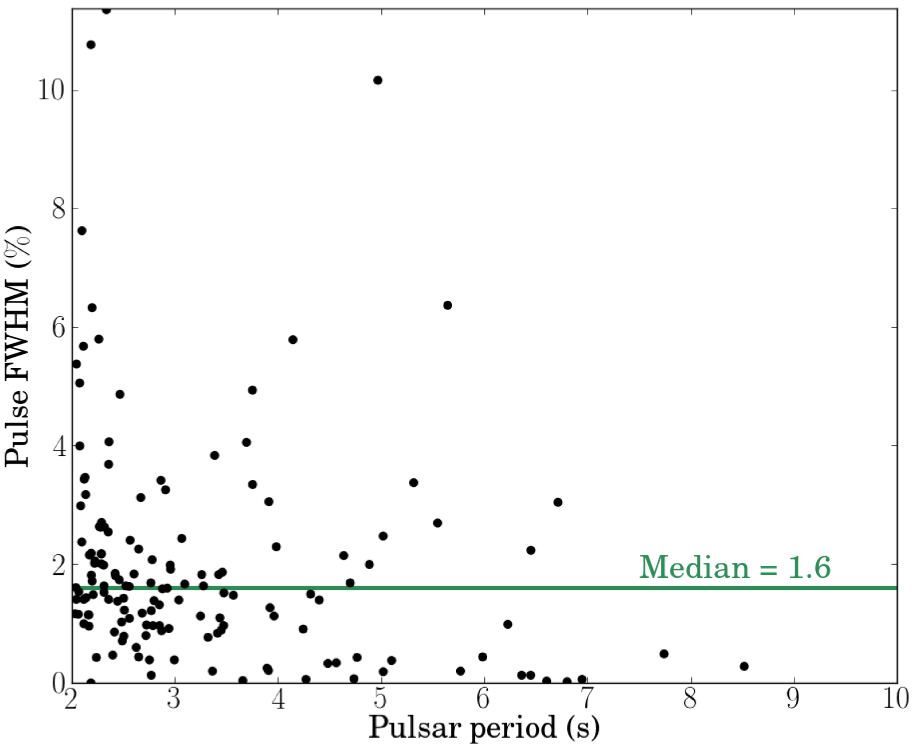


Figure 3.4: Distribution of known pulse FWHM[4] (in units of $\%$ of the period) of radio pulsars. The median of the distribution is represented by the green solid line. These measurements were made at an observing frequency of $\sim$ 1.4 GHz.

3.3 Implementation of the Fast-Folding Algorithm in PALFA

Parts of the Fast-Folding Algorithm code are taken from a github FFA package[5], written as a python and C program by Erik Petigura (Hubble Fellow at the California Institute of Technology), and developed for transient searches in optical data taken by NASA's Kepler mission (Petigura et al., 2013) which searches for exoplanets . More specifically, the parts of the code that wrap the time series, pad it, perform the folding and the summations were taken from Petigura's work.

---

[5] Available at https://github.com/petigura/FFA

For this MSc thesis, I have written a python program that implements the FFA periodicity search into the PALFA analysis software. It includes a detrending procedure aimed at reducing the effect of red noise on the input time series. This is done by applying a running median filter, where the size of the filtering window is set to 50 sec (using smaller filtering windows could potentially remove power from long-period pulsars if present in the time series). The data are then normalised with respect to the maximum value of the time series. The program also performs an initial downsampling of the data to increase the speed of the FFA search. Subsequent dynamical rebinning routines are applied in the code to search for multiple pulse duty-cycles (more details can be found below). Signal-to-noise ratio calculations, candidate selection and sifting procedures are also incorporated in this program. Periodograms similar to Figure 3.3 can also be generated by the program.

When implementing the FFA on pulsar searches, one has to choose which pulsar parameters to explore while considering the computation time allowed. The pulsar parameters that we consider in the PALFA search pipeline are the following:

1. **The period**

   The trial periods range from a minimum period of 500 ms to a maximum of 30 seconds. Even though the Fast-Folding Algorithm is designed to be fast, it is still computationally expensive to apply it for short periods since those are the ones that result in a large number of profiles to be statistically evaluated. This is one of the reasons why the blind period search is restricted to periods longer than 500 ms. Moreover, 500 ms is approximately the period at which one notices a decrease in the sensitivity of PALFA at small dispersion measures, as shown in Lazarus et al. (2015; see Chapter 4). It is not worthwhile to look for periods larger than 30 sec as the PALFA observations are only $\sim$268 sec: folding fewer than $\sim$ 10 pulses will generally not result in significant signals. It is hoped that the single pulse search done in the PALFA pipeline can potentially identify pulses from such slow pulsars.

2. **The pulse width**

   To explore the pulse width parameter, we perform rebinning many times in the FFA process such that the sampling interval ranges from $\sim$2 ms to a few seconds, depending on the trial period and the pulse duty-cycle we are searching for at each step of the process. An initial rebinning of the time series is performed such that the sampling interval goes from 65.5 μs (the default time resolution of PALFA search observations) to about 2 ms, which is necessary to have acceptable computing times when looking for periods larger than 500 ms. Then, an extra rebinning that depends on the trial period is done such that the minimum duty-cycle is around $\sim$ 0.4%. To do so, we split the 500 ms-30 sec trial period range into 6 sub-ranges of periods such that all trial periods in this range will use a time series containing the same number of samples. Finally, for each trial period, we search for pulsars with different pulse widths by doing additional downsampling by various factors, at different phases. We are then sensitive to pulse duty-cycles ranging from a minimum of $\sim$0.4% up to a maximum of nearly $\sim$25%.

Re-binning is chosen over convolutions with a boxcar function to solve for the pulse width of the signal to reduce the computation time, which is crucial considering the parameter space we are exploring. Moreover, lowering time resolution speeds up the computation time for the FFA as the sampling interval increases, since the number of profiles to be examined decreases with the number of samples in the time series.

Figure 3.5 shows how the rebinning of the time series is performed. The Fast-Folding Algorithm is applied on each of the coloured boxes in the Figure, meaning that the FFA is applied on a total of 54 slightly different (downsampled at different phases, see Figure 3.5) time series. Starting from an initial sampling interval Δt, the time series is downsampled consecutively by factors of 2^1, 2^2 and 2^3, and then by factors of 3^1, 3^2 and 3^3, always rebinning at different phases. For example, for the shortest period range we are searching over (500 ms to 1 seconds), the initial sampling interval is 2 ms, implying that the trial pulse widths are approximately 2 ms, 4 ms, 6 ms, 8 ms, 16 ms, 18 ms and 54 ms.

3. **The DM**

   Since the DMs of the pulsars to be discovered are yet unknown, a large number of DM trials must be used in the search. We search only from DM = 0 pc cm^{-3} to DM $\cong$ 3000 pc cm^{-3} in steps of 5 pc cm^{-3}, meaning that approximately 600 dedispersed time series are searched with the FFA. We are not going beyond DM values of 3000 pc cm^{-3}, which is slightly more than the maximum dispersion measure expected in the regions of our Galaxy that PALFA surveys ($\sim$ 1350 pc cm^{-3}) according to the NE2001 model (Cordes & Lazio 2002). The probability of finding normal pulsars with flux densities of a few mJy outside our Galaxy is quite low as pulsars are weak radio sources, and this is why searching beyond our Galaxy is not worth the large increase in computation time that a search over higher DM values would require.

Because the FFA is optimal for long-period pulsars, the pulse widths in milliseconds are expected to be quite large: we do not expect sub-millisecond pulse widths. This implies that it is not necessary to go through fine DM steps in order to detect pulsars. This is necessary for rapidly rotating pulsars because the pulses can be dispersed within frequency channels, but not for slow pulsars (i.e., the pulse widths are approximately equal or larger to the dispersion of the pulses within channels). The DM step size was chosen such that the amount of processing is minimized while avoiding sensitivity loss from channel smearing due to dispersion.

3.3.1 Candidate selection

For each dedispersed time series analysed by the FFA, a list of final period candidates is produced. Any period with a signal-to-noise ratio higher than 5 is included in this list. A simple signal-to-noise ratio metric is used while evaluating the significance of the generated profiles, which assumes that the profile has one single peaked pulse, that this pulse is constant in phase and that the pulse profile power is captured within a single bin (i.e., the pulse width and the sampling time interval are similar). The metric is as follows:

$$S/N_{\text{profile}} = \frac{\text{Max}_{\text{profile}} - \text{Median}_{\text{profile}}}{\sigma_T} \tag{3.2}$$

where the standard deviation σ_T of the profile accounts for downsampling and the number of periods in the time series (M):

$$\sigma_T = \sigma_{\text{data}} * \sqrt{\frac{N}{P} - z} \tag{3.3}$$

where σ_{data} is the standard deviation of the decimated data before entering the FFA and z is the fraction of padded data (i.e., the number of profiles that must be padded such that M is a power of 2).

Reducing computation time is one of the reasons why this simple metric was chosen over other possibilities, such as one where the standard deviation would be calculated only over a certain portion of the profile, when the pulse is off, for example. Such a metric can be effective where the periodicity search is limited and applied to a small number of dedispersed time series. Kondratiev *et al.*,(2009) conducted an FFA search with such a metric to obtain a numerical score of the folded profiles. Another option was to fit a polynomial to the profiles and calculate the χ^2 value, such that we are sensitive to profiles containing more than one pulse and/or to multiple phase windows. Again, this method requires too much computation power to be applied to a large-scale pulsar search.

The period, the S/N and the sampling interval are recorded for every candidate. The list of candidates is then sifted in order to remove duplicate and harmonically related periods (see §3.3.2).

Once all the possible dispersion measures ($0 \text{ pc cm}^{-3} < \text{DM} < 3000 \text{ pc cm}^{-3}$), period sub-ranges and pulse widths have been searched and the candidates sifted, only the candidates with S/N > 6 are folded. Figure 3.6 shows the distribution of signal-to-noise ratios one obtains from the FFA program when applied to a time series containing no periodic signal and where the noise is normally distributed. This limit is applied to reduce the number of false positives. The candidates are then uploaded to a PALFA's online Candidate Viewer application for human inspection and classification.

3.3.2 Sifting

For each dedispersed time series, many hundreds of FFA candidate periods are selected and then recorded in a final candidate list based on their signal-to-noise ratio. It is necessary to remove irrelevant or redundant (i.e. the same signal seen in multiple DM trials, or harmonics of the same signal) candidates in the final candidate list via sifting to reduce the number of candidates that will eventually be inspected. This important routine is also used for typical periodicity candidates produced by the Fast Fourier analysis, as part of the data analysis pipeline.

The first step of the sifting procedure consists of removing candidates with similar periods found at multiple DM values (with possibly different pulse widths), keeping only the candidates with the highest signal-to-noise ratio. Then, candidates showing some issues regarding their dispersion measure are excluded from the final list: the candidates found at only one DM are rejected as those are typically false positives, as well as those with a maximum DM smaller than 2 pc cm^{-3}, as those are the ones that are likely RFI and not astrophysical. Finally, all lower-significance candidates that are harmonically related are rejected. The number of other candidates with related periods that were found at multiple DM values is noted for each final candidate, as this information is useful when trying to identify RFI, which tends to appear very often at multiple, low DM values.

Using these criteria, the list of final candidates goes from a few thousand candidates to less than on average fifty candidates per PALFA beam.

3.4 Testing on Artificial Pulsars Injected in Gaussian Noise

To confirm and quantify the ability of the algorithm to recover pulsar signals, artificial signals were injected in Gaussian noise using the `SIGPROC`[6] program `fake`. Spin periods of 2 sec, 5 sec, 10 sec, 14 sec and 20 sec were used for this analysis, with a narrow, fixed pulse FWHM of 1.6 $\%$ (i.e., the median value of pulse duty-cycle for long-period radio pulsars for which the pulse width is measured). To match real PALFA data, the sampling interval of the fake observation was set to 65 μs with a total integration time of 268 sec. An example of a periodicity plot generated by `PRESTO`'s `prepfold` command for one of these constructed pulsars can be found in Figure 3.7.

The five generated pulsars were processed through the FFA program as well as through `SIGPROC`'s `seek` and `PRESTO`'s `accelsearch` programs, two Fourier-based periodicity searches (both including incoherent harmonic summing up to 16 harmonics). Figure 3.8 shows the periodogram

---

[6] https://github.com/SixByNine/sigproc

resulting from the FFA search for the artificial pulsar represented in Figure 3.7. The signal-to-noise ratios obtained from the 3 periodicity searches for the five injected signals are plotted in Figure 3.9. The S/N of a candidate produced by a time-domain search is based on a folded profile, while the S/N of a FFT candidate is based on the power contained in Fourier bins in a power spectrum. Therefore, one has to be careful when comparing directly the signal-to-noise ratios of a candidate period from the two search types because of the different nature of the statistical significance of a candidate (i.e., a S/N of 20 from a FFA search should not be considered as 2 times stronger than a FFT candidate with a S/N of 10).

Figure 3.9 suggests that, even in white noise, the signal-to-noise ratio of the detections decrease with increasing period. This is expected since there are fewer pulses in the time series when injecting longer periods in an observation with a fixed integration time. However, when looking at a period of 20 sec, the S/N from the FFA remains an order of magnitude larger than the S/N folding threshold set in the PALFA analysis pipeline (S/N must be greater than 6 for a periodicity candidate to be folded, see Chapter 2), while the resulting S/N from the frequency-domain searches is less than a factor of $\sim$2 above the folding threshold. When comparing the significance of the detections at 2 sec and 20 sec, the signal-to-noise ratio produced by the FFA dropped by only a factor of 1.5, whereas it dropped by factors of 4 and 3.1 for the `SIGPROC`'s FFT and the `PRESTO`'s FFT searches, respectively. Moreover, some of the Fourier-based detections were identified via their harmonics (see Table 3.1), while the FFA search always identified the fundamental period as being the strongest signal. This highlights the advantage of coherent harmonic summing employed by the time-domain search.

In summary, this analysis demonstrated both the ability of the FFA to recover pulsars and its advantage over frequency-domain searches in the long-period regime when in the presence of ideal, Gaussian distributed noise. A similar study is presented in Chapter 4, where artificial pulsar signals have been injected in real observational data rather than in white noise to examine the response of the Fast-Folding Algorithm when searching for pulsars in a large-scale survey.

Table 3.1: Signal-to-noise ratio of 3 periodicity searches

	FFA		**`seek`**		**`accelsearch`**	
Injected Period (s)	Detected Period (s)	S/N	Detected Period (s)	S/N	Detected Period (s)	S/N
2.00	2.00	69.6	2.00	45.6	2.00	29.9
5.00	5.00	65.8	2.50◇	31.9	5.00	25.4
10.00	10.00	58.2	5.00◇	30.4	5.00◇	21.9
14.00	14.00	51.5	6.99◇	23.3	7.00◇	16.4
20.00	20.00	46.6	6.67◇	10.8	6.67◇	9.7

◇ Harmonic detections

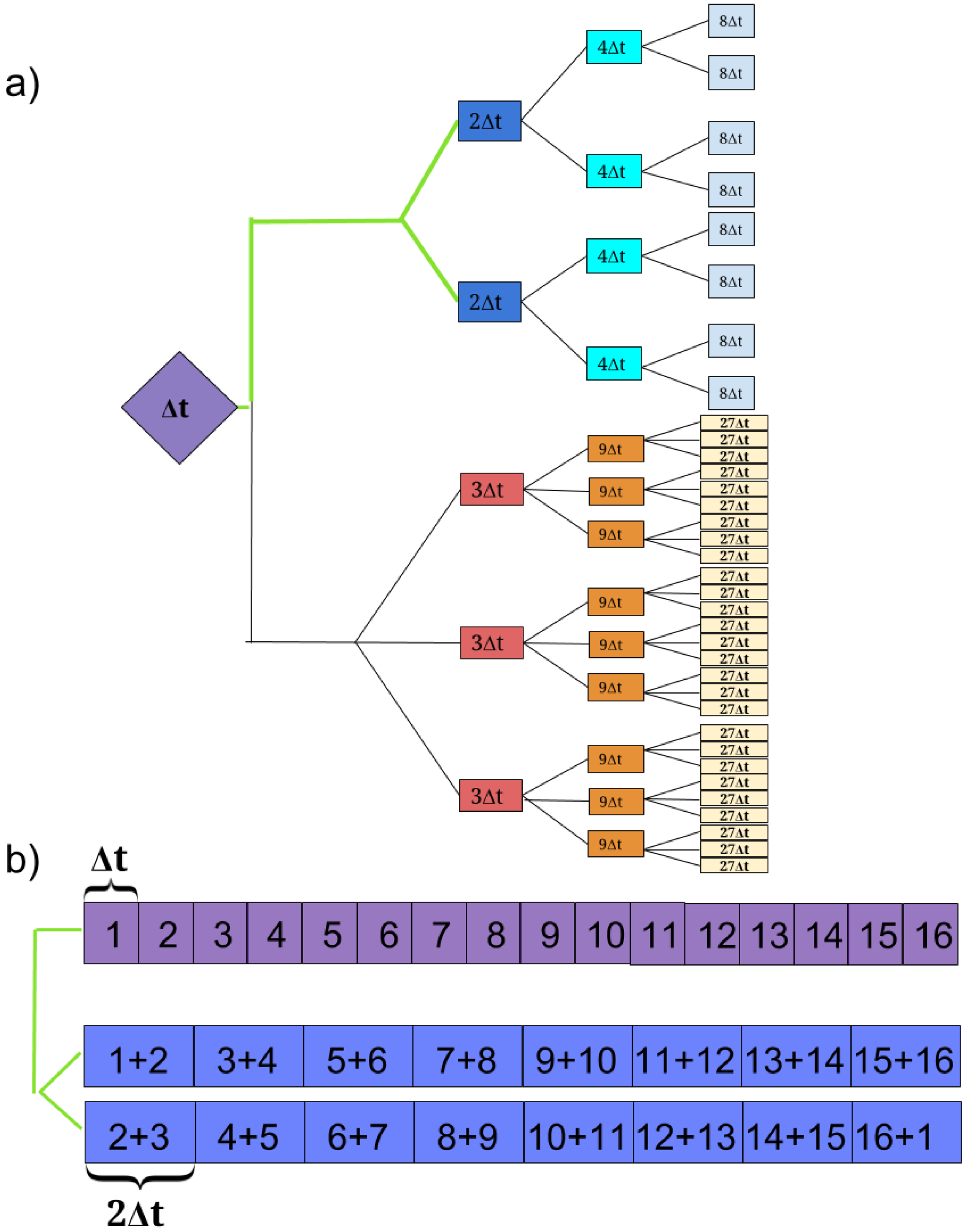


Figure 3.5: Tree diagram showing the different pulse widths that are searched for in PALFA's implementation of the Fast-Folding Algorithm (for a given period sub-range). Rebinning of the time series by factors of 2, 3, 4, 8, 9 and 27 is performed at different phases. For each period sub-range, a total of 54 different time series are analysed by the FFA.

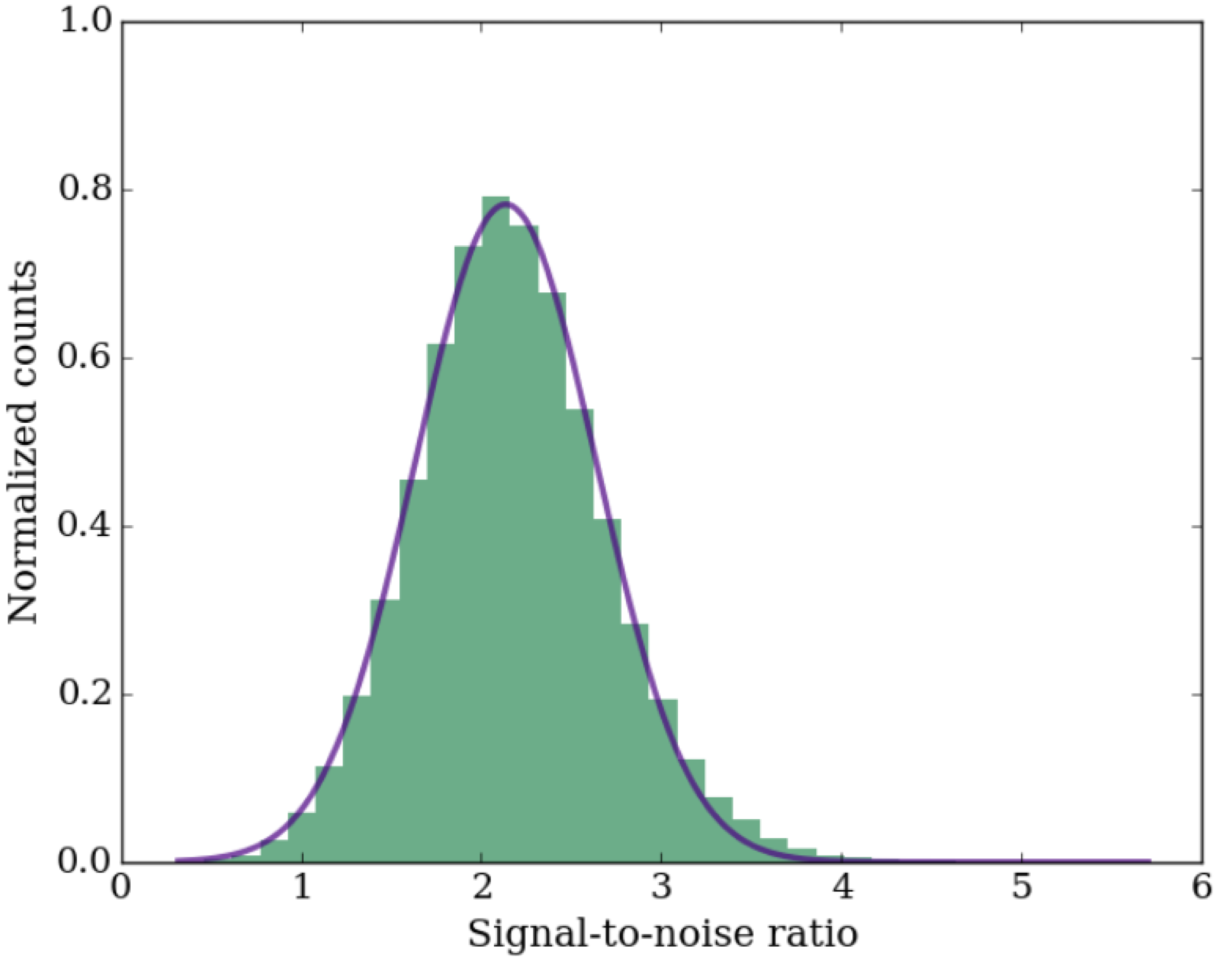
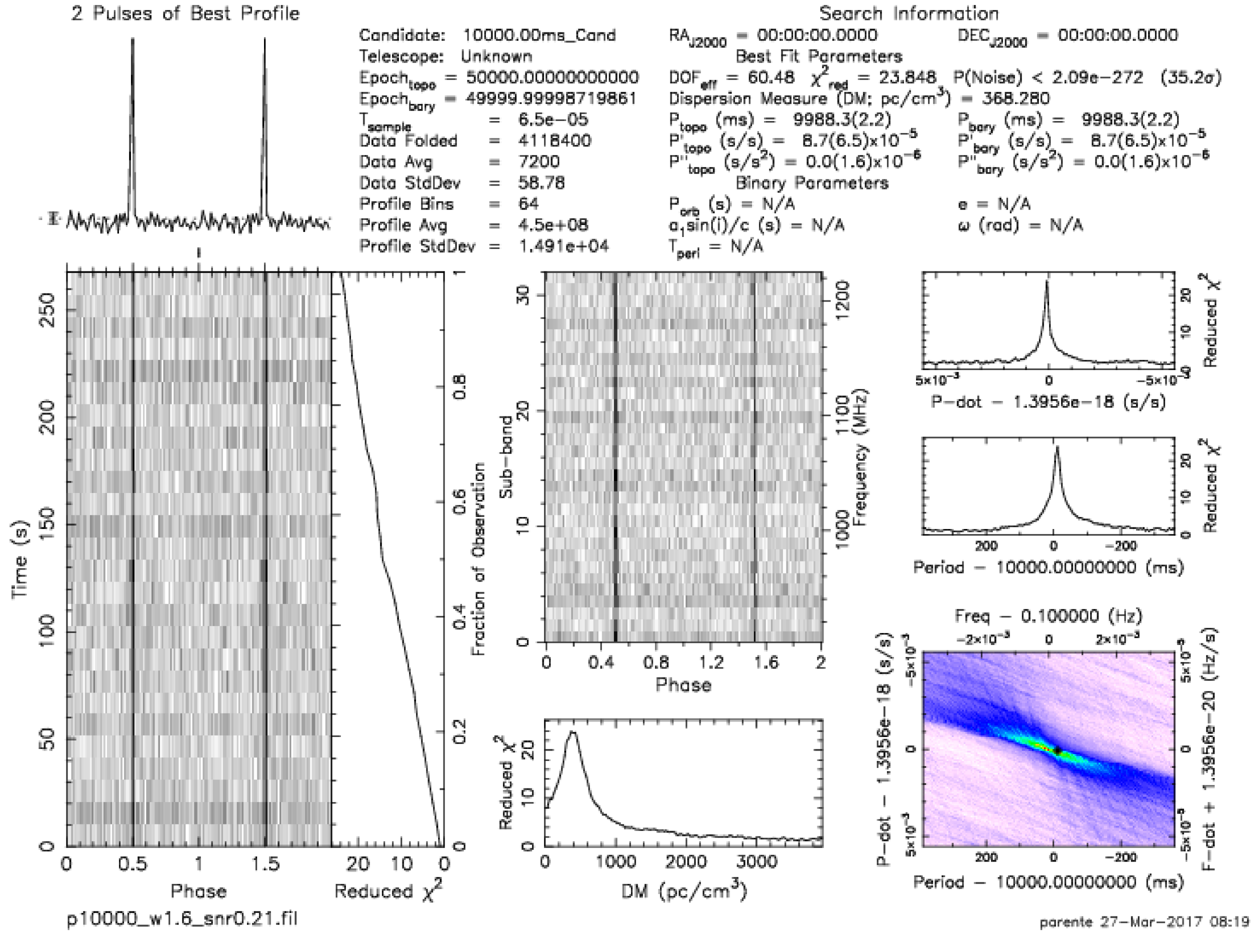


Figure 3.6: Distribution of signal-to-noise ratio of the profiles evaluated when searching in normally distributed noise containing no periodic signals with the implementation of the Fast-Folding Algorithm presented in this thesis. The S/N values follow a Gaussian distribution (purple curve) with a mean value of 2.1 and a standard deviation of 0.5. Considering the metric used to evaluate pulse profiles (Eq. 3.2), we do not expect the Gaussian distribution to be centered at zero, as the maximum value of the profile will always be greater than the mean value of the profile for time series where the intensity is not flat across the samples.



```
Candidate:  10000.00ms_Cand          RA_J2000 = 00:00:00.0000            DEC_J2000 = 00:00:00.0000
Telescope:  Unknown                           Best Fit Parameters
Epoch_topo = 50000.00000000000       DOF_eff = 60.48  χ²_red = 23.848  P(Noise) < 2.09e-272  (35.2σ)
Epoch_bary = 49999.99998719861       Dispersion Measure (DM; pc/cm³) = 368.280
T_sample          =  6.5e-05         P_topo (ms) =  9988.3(2.2)          P_bary (ms) =  9988.3(2.2)
Data Folded       =  4118400         P'_topo (s/s) =  8.7(6.5)x10^-5     P'_bary (s/s) =  8.7(6.5)x10^-5
Data Avg          =  7200            P''_topo (s/s²) = 0.0(1.6)x10^-6    P''_bary (s/s²) = 0.0(1.6)x10^-6
Data StdDev       =  58.78                    Binary Parameters
Profile Bins      =  64              P_orb (s) = N/A                     e = N/A
Profile Avg       =  4.5e+08         a_1sin(i)/c (s) = N/A               ω (rad) = N/A
Profile StdDev    =  1.491e+04       T_peri = N/A
```

Figure 3.7: Periodicity plot generated by `PRESTO`'s `prepfold` command of an artificial 10 sec period pulsar injected in Gaussian noise, generated with the `SIGPROC` program `fake`. This signal was dispersed at a DM of 400 pc cm^{-3}. See Figure 2.6 for the description of the different panels in this plot.

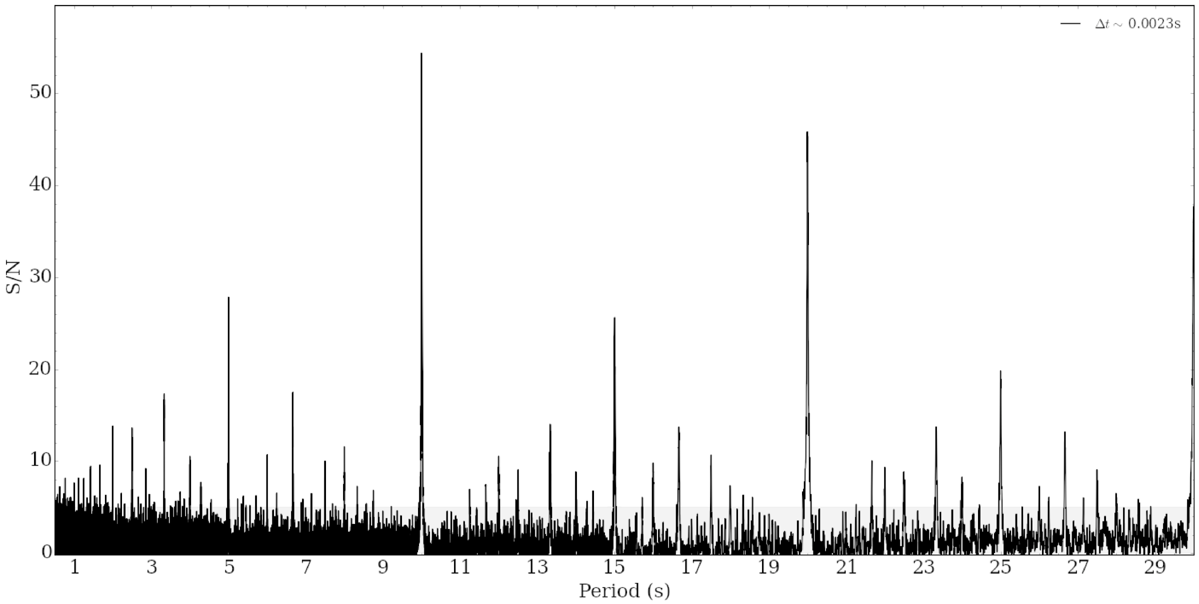


Figure 3.8: Periodogram generated by the FFA program for the 10 sec period artificial pulsar shown in Figure 3.7. Note the maximum S/N at the fundamental period and the harmonic structure. The initial sampling interval used to produce this periodogram is 1.9 ms (dynamic downsampling is performed, according to the period searched).

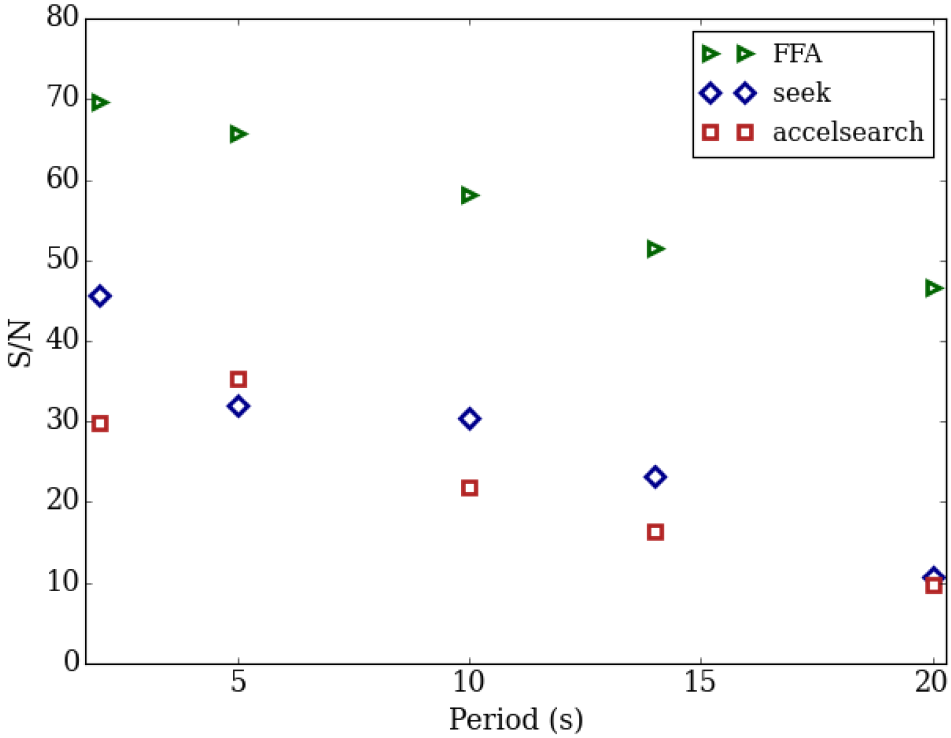


Figure 3.9: Results from the analysis where artificial pulsar signals are injected in normally distributed noise. The plot shows the S/N of the detections for the FFA (solid line), `SIGPROC` program `seek` (dashed line) and `PRESTO` program `accelsearch` (dotted line) as a function of injected period. Note the more important degradation in S/N when searching over longer periods for the two frequency-domain searches, as compared to the FFA search.

Chapter 4

Results

Multiple factors, from both the instruments and the targeted astrophysical sources, contribute to the ability of a large-scale survey to discover pulsars. For any survey, the minimum detectable phase-averaged flux density of an observed radio pulsar depends on parameters such as the period, the pulse width and the dispersion measure of the pulsar, the background sky temperature, the gain of the antenna and the receiver noise temperature. The radiometer equation (Dewey *et al.*, 1985) is typically used to estimate the mean flux density of pulsars (usually calibrated in Jy) :

$$S_{mean} = \frac{(S/N)_T \Big(T_r + T_{sky}(l,b)\Big)}{G\sqrt{N_p t_{obs} \Delta f}} \sqrt{\frac{W}{P-W}} \tag{4.1}$$

where $(S/N)_T$ is the time-domain signal-to-noise ratio of the pulse, T_r is the receiver system noise temperature, $T_{sky}(l,b)$ is the sky temperature at the Galactic coordinate (l,b), G is the antenna gain (which depends on the angular offset of the source from the centre of the beam and on the zenith angle), N_p is the number of polarization channels, t_{obs} is the integration time of the observation and Δf is the total observing frequency bandwidth. The pulse width, W, and the period of the pulsar, P, also affect the mean flux density: long-period pulsars with narrow integrated pulse profiles have lower S_{mean}.

This expression, however, does not include the effect of red noise and RFI, which varies according to the environment at the telescope. A more complete analysis is required to assess the true sensitivity of the survey (i.e., to establish the minimum detectable mean flux density, S_{min}), which depends on a chosen integrated pulse signal-to-noise threshold. This type of analysis has been carried through in L15 to evaluate the sensitivity of the PALFA survey by injecting synthetic pulsar signals into real observation data, free of astrophysical signals. This work was done using the previous version of the survey (PALFA3), which has been recently updated to the new version (PALFA4) described in Chapter 2.

§4.1 explains how artificial pulsars are constructed, and §4.2 summarizes the results from the analysis conducted in L15. The same synthetic signals were then used to evaluate the ability of the Fast-Folding Algorithm to recover pulsar signals so that we can estimate the improvement in the survey sensitivity when using this time-domain technique in long-period pulsar searches. The results from this analysis are presented in §4.3.

4.1 Injecting Synthetic Pulsar Signals in Real Observational Data

To assess the PALFA survey sensitivity, artificial pulsar signals were constructed and injected in real survey data using the program `injectpsr`[1], written by Patrick Lazarus (L15). This program generates smeared, scattered and scaled pulse profiles which are then added to the data at regular time intervals corresponding to the chosen period. To scale the profiles properly, observations of the radio galaxy 3C 138 were carried out in December 2013. During these observations, a calibrating noise diode was turned on so that the flux density of the diode could be compared to that of the galaxy (which has a known flux density from the literature). Per-channel scaling factors between flux density and the observation data units were then calculated (see L15 for more details). These scaling factors are used to obtain the targeted phase-averaged flux density of the artificial pulsar signals.

When constructing a signal with `injectpsr`, the pulse profiles generated have single von Mises components, a periodic analog to a Gaussian distribution that is often used to model pulse profiles of pulsars, whose probability density function is expressed as follows:

$$f(x, \mu, \kappa) = \frac{e^{\kappa \cos(x-\mu)}}{2\pi I_0(\kappa)} \tag{4.2}$$

---

[1] Available at https://github.com/plazar/presto

where μ is the mean of the distribution, κ is a reciprocal measure of dispersion (analog to 1/σ^2 in a Gaussian distribution), and $I_0(\kappa)$ is the modified Bessel function of order 0. `injectpsr` allows the user to specify the FWHM of the function. Dispersive smearing and scattering are then applied to the profile, where the amount of broadening caused by scattering is determined by the specified DM and the observing frequency, ν, according to the following relation (Bhat *et al.*, 2004):

$$\log \tau_{scatt} = -6.46 + 0.154 \log DM + 1.07(\log DM)^2 - 3.86 \log \nu. \tag{4.3}$$

Figure 4.1 illustrates an example of an artificial pulsar where the frequency channels have been scaled using the calibration procedure described previously. The program then records the modified data into a `SIGPROC` "filterbank" file format. We can then use `injectpsr` to create synthetic pulsars for which we chose S_{mean} and then characterise the sensitivity of a survey at each point in (period, DM, pulse FWHM) phase-space.

4.2 The Sensitivity of PALFA3

A realistic sensitivity analysis of the `PRESTO`-based PALFA pipeline (previous version, PALFA3) was conducted in L15 to evaluate the performance of the PALFA survey at finding different types of pulsars using `PRESTO`'s `accelsearch` program, which searches for periodic signals in the frequency-domain.

To properly evaluate the dependence of the survey sensitivity on the period, the DM and the pulse duty-cycle of pulsars, artificial pulsar signals were constructed with `injectpsr` and then injected in 12 PALFA observations (the data were collected in 2013). The selected observations were initially searched with the PALFA processing pipeline and no astrophysical signals were detected. Those observations contain a variety of types of radio frequency interference, which is important to assess the true sensitivity of the survey. A total of 17 different trial periods were chosen, ranging between less than 1 ms up to nearly 11 sec, along with 6 DM trial values, ranging from 10 pc cm^{-3} to 1000 pc cm^{-3}. Single centred von Mise components with FWHM between 1.5$\%$ and 24.3 $\%$ were used to generate the signals. Injecting all possible combinations would have resulted in too large a computational burden. Therefore, only the pulses with FWHM of 2.6 $\%$ were injected in all data files; the remaining trial pulse FWHM were injected in only one of the

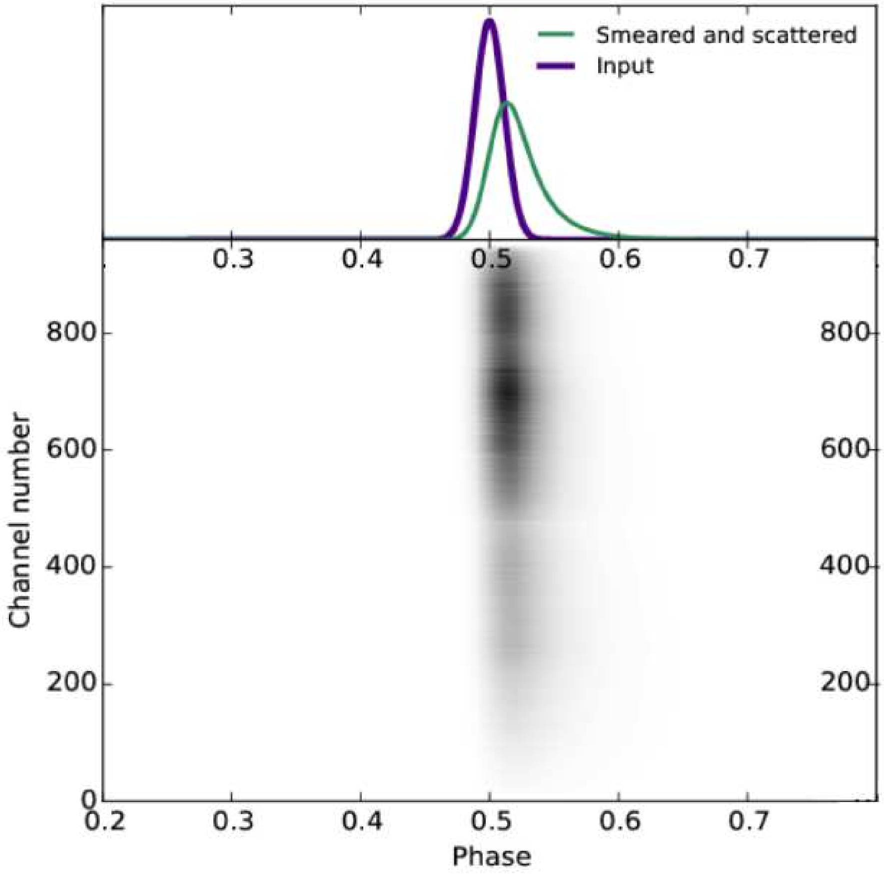


Figure 4.1: *Top panel*: The purple curve is the pulse profile of a 50 ms period synthetic pulsar consisting of a single centred von Mises component (FWHM = 2.6 %), generated by `injectpsr`. The green curve is the same pulse profile once scattering and smearing (DM of 400 pc cm^{-3}) have been applied. *Bottom panel*: Intensity of the broadened signal as a function of phase and frequency channel number (960 channels in total, from 1225 MHz to 1525 MHz) once the channels have been scaled according to the calibration procedure.

files. Table 4.1 lists the pulsar parameters used in this analysis.

Once the injection of a trial set of parameters was completed, the generated filterbank file was fully processed with the PALFA's analysis pipeline. Candidate periods were then produced by the `accelsearch` program (which incoherently summed up to 16 harmonics). This was done for all combinations of trial parameters, and each of them was then evaluated to determine whether the injected signal was recovered by the periodicity search or not. The mean flux density was then

Table 4.1: Parameters of the synthetic pulsar signals used in the Fourier Transform-based analysis

Parameter	Possible Values						
Period (ms)	0.77	1.10	2.22	5.22	10.87	18.51	26.97
	61.63	126.18	286.56	533.32	850.16	1657.50	2643.41
	3927.01	5581.90	10965.53	...	...	...	...
DM (pc cm^{-3})	10	40	150	325	400	600	1000
FWHM (% Phase)	1.5	2.6	5.9	11.9	24.3	...	...

adjusted, depending on whether the artificial pulsar was recovered or not (i.e., if the signal was not detected/detected, the flux was increased/reduced until detection/non-detection), until the minimum detectable phase-averaged flux density (S_{min}) was established.

An important result from this analysis is that there is a clear mismatch between sensitivity curves measured (solid curves in Figure 4.2) and the ideal case scenarios, predicted by 2 slightly different versions of the radiometer equation (dashed curves in Figure 4.1) when searching in the long-period regime. This suggests that the survey is not as sensitive as it should be. Indeed, this work showed that a reduction in the survey sensitivity is noticeable for spin periods as short as $\sim$100 ms (at low dispersion measures). The radiometer equation predicts a constant minimum detectable flux density of $\sim$15 μJy when searching for periods larger than $\sim$1 sec, but we can clearly see that S_{min} increases toward longer periods. At a period of $\sim$11 sec, the measured S_{min} are 10 to 20 times larger than the predicted value, for dispersion measures of 10 pc cm^{-3} and 600 pc cm^{-3}, respectively (thick solid lines in Figure 4.2).

In this work, synthetic pulsar signals were also injected in Gaussian noise (thin solid lines in Figure 4.2). These curves also rise at longer periods, although not as rapidly as the curves from the real data injections. This could suggest that RFI and red noise alone cannot explain the mismatch between the measurements and the predictions.

4.3 The Sensitivity of the Fast-Folding Algorithm

The analysis described in §4.2 was reproduced to assess the sensitivity of the Fast-Folding Algorithm in the long-period parameter space. In the work presented here, only pulsar signals with

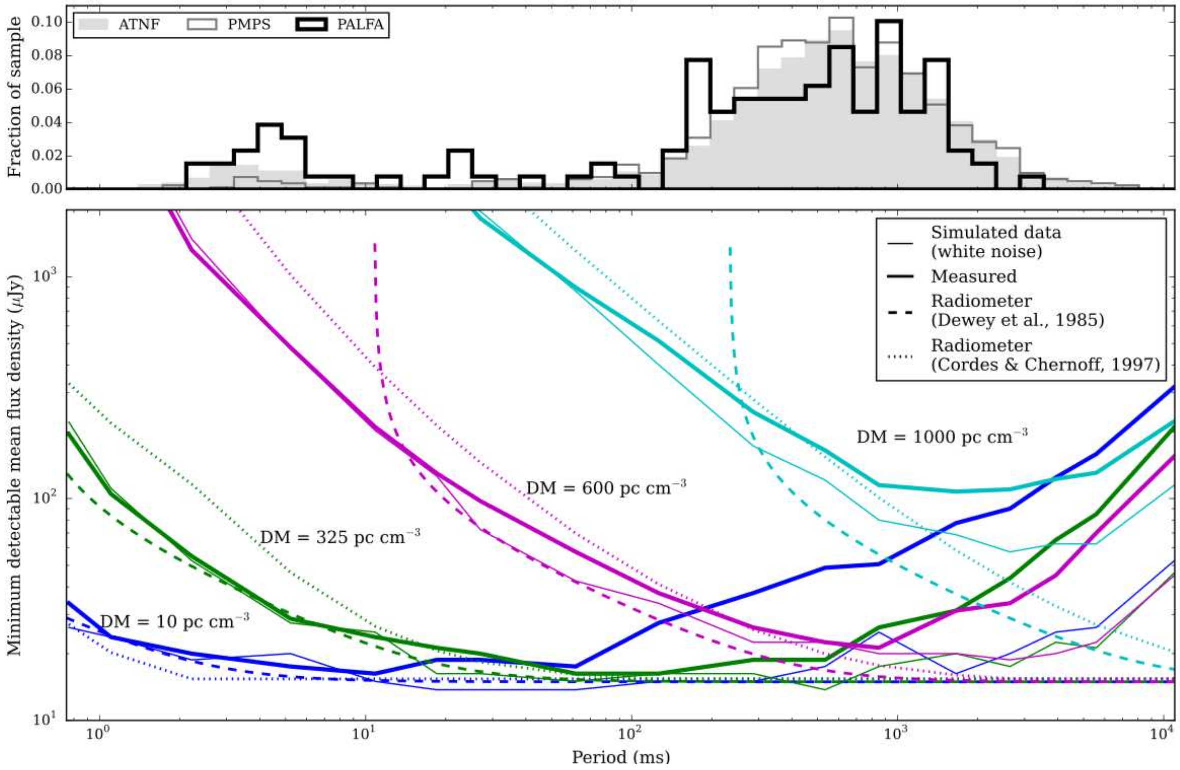


Figure 4.2: *Top panel*: Period distribution of radio pulsars, as listed in the ATNF pulsar catalogue (http://www.atnf.csiro.au/research/pulsar/psrcat/). This includes discoveries from the PALFA survey (as of 2015) and from the Parkes Multibeam Pulsar Survey (PMPS). *Bottom panel*: Minimum detectable average flux density (S_{min}) curves showing the sensitivity of PALFA (solid lines) as measured using synthetic signals with pulse duty-cycle of $\sim$2.6 % injected in real survey data. The thin solid lines are the measured curves when the signals are injected in white noise. The S_{min} values reported are the median value of the minimum detected mean flux density from the injections in the 12 observation data files. The dashed and the dotted lines represent the predicted survey sensitivity from 2 slightly different versions of the radiometer equation (see Dewey *et al.*, 1985 and Cordes & Chernoff, 1997). One can see a clear mismatch between the measured minimum detectable flux and the predicted flux in the long-period regime. Figure taken from L15.

periods longer than 500 ms were injected into the same real survey data that were used in L15, and for consistency, the same values of dispersion measure were used. However, we have also injected signals with periods of $\sim$15 and $\sim$21 sec, rather than stopping at 11 sec, so that we have a better understanding of the responsiveness of the FFA in the long-period regime. Synthetic pulsars with FWHM pulses of 0.5 %, 1.5 %, 2.6 %, 5.9 % and 11.9 % were injected into the PALFA observations, and as in L15, only the signals with a pulse FWHM of 2.6 % were injected in all 12 survey data files: the other values of pulse FWHM were injected in only one of the data files (the same as the one used in L15). The list of pulsar parameters used in this analysis can be found in Table 4.2.

Figure 4.3 shows the median of the minimum detectable mean flux density required for the FFA to detect the injected signals having a FWHM of $\sim$2.6 $\%$. The curves obtained from this study of the FFA (solid lines) do not flatten out at longer periods, as opposed to what the radiometer equation predicts. However, they do not rise as rapidly as the curves from the Fourier-based search technique reported in L15, meaning that we observe a gain in sensitivity. The efficiency of the FFA search does not seem to vary strongly with DM at the longest periods, while we can see that an FFT can detect more easily high DM pulsars compared to low DM ones. This can be explained by the fact that the RFI-excision routines conducted in the processing pipeline tend to remove signals that are found at low values of DM, which are often flagged as interference.

Figure 4.4 illustrates the improvement factor of the FFA-based search over the FFT-based search as a function of period for the 4 trial DM values used, where the pulse FWHM = 2.6 $\%$ is kept fixed (see Table 4.3). As expected, the gain in sensitivity is greater as we go toward longer periods and toward smaller dispersion measures (as there is a reduction in the efficiency of the FFT-based search at low DMs).

The second part of the analysis consisted in injecting signals with different pulse duty-cycles into one of the observational data files. One can see from the results (see Figure 4.5) that, for all DMs and pulse FWHM at longer periods, the FFA curves (solid lines) rise more slowly than the ones from the FFT, implying that the time-domain search is more efficient than the frequency-domain search in this regime. The gain in sensitivity is unexpectedly greater when considering broad pulses (see Figure 4.6 and Table 4.4). This is surprising because the advantage of a time-domain search over a frequency-domain search is that the coherent summing of all harmonics makes the search especially responsive to narrow pulses. Even if the factor of improvement in the sensitivity is more significant for broader pulses, Figure 4.6 shows that the FFA is also capable of detecting weak signals with narrow pulses missed by `PRESTO`'s `accelsearch` program, in the low modulation frequency regime.

Table 4.2: Parameters of synthetic pulsar signals used in the FFA analysis

Parameter	Possible Values				
Period (ms)	533.3	1657.5	3927.0	5581.9	10965.5
	14965.5*	21427.7*	...	...	
DM (pc cm^{-3})	10	150	400	600	...
FWHM (% Phase)	0.5*	1.5	2.6	5.9	11.9

* New trial elements, not included in the sensitivity analysis conducted in L15.

Table 4.3: Improvement factors from the FFA search at a pulse FWHM = 2.6 %

DM (pc cm^{-3})	Period (ms)	Improvement Factor
10	533.3	1.0 ±0.1
	1657.5	1.5 ±0.1
	3927.0	1.4 ±0.1
	5580.9	1.9 ±0.2
	10964.5	2.7 ±0.2
150	533.3	0.9 ±0.1
	1657.5	1.1 ±0.1
	3927.0	1.2 ±0.3
	5580.9	1.6 ±0.1
	10964.5	2.3 ±0.2
400	533.3	0.8 ±0.2
	1657.5	1.0 ±0.1
	3927.0	1.0 ±0.1
	5580.9	1.4 ±0.1
	10964.5	2.0 ±0.1
600	533.3	0.7 ±0.5
	1657.5	0.9 ±0.5
	3927.0	0.9 ±0.1
	5580.9	1.1 ±0.1
	10964.5	1.5 ±0.1

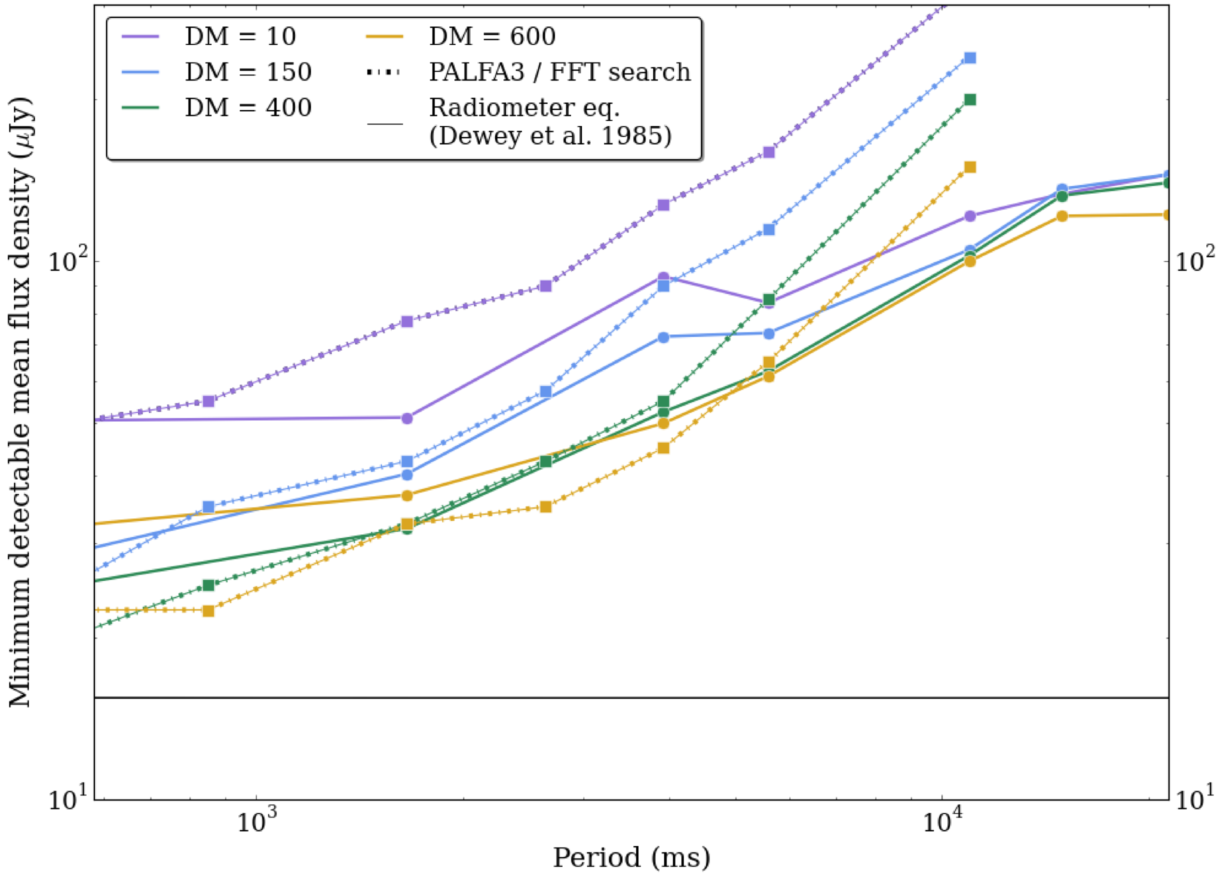


Figure 4.3: Minimum detectable average flux density (S_{min}) curves showing the sensitivity of the FFA (solid lines) as measured using synthetic signals with pulse duty-cycle of 2.6 $\%$ injected in real survey data. The thin black line represents the survey sensitivity predicted by the radiometer equation (Dewey *et al.*, 1985), while the dashed lines are the S_{min} calculated from the Fourier-based search (from L15). The S_{min} values reported are the median value of the minimum detected mean flux density from the injections in the 12 observation data files. Note: the lines simply connect the data points.

4.4 Discussion

Despite the noticeable improvement of the survey sensitivity of PALFA when using the time-domain search to look for long-period pulsar signals in real observational data, both the Fourier-based and FFA-based search types are not able to recover weak signals that should be detectable according to the radiometer equation predictions (e.g. Dewey *et al.*, 1985). Even if the FFA search technique

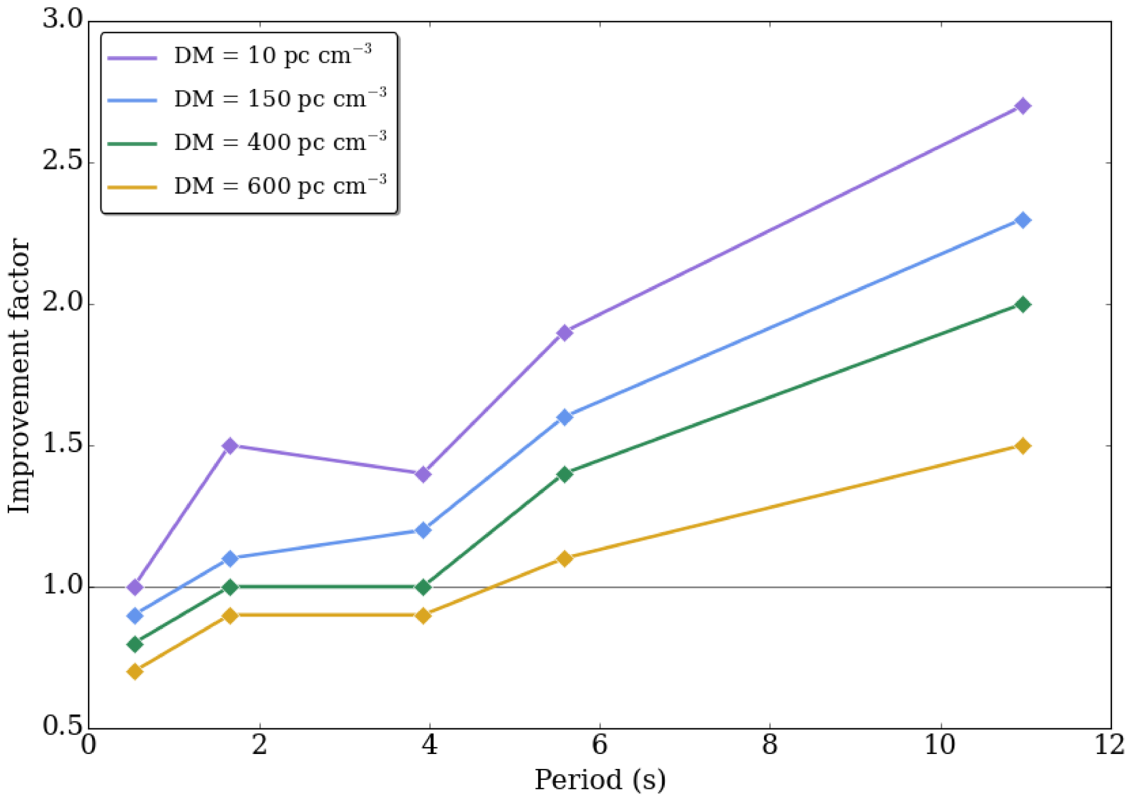


Figure 4.4: Factors of improvement in the sensitivity of the FFA-based search over the FFT-based search as a function of period, for different DM values. The sensitivity of the FFA is greater than that of the FFT when the curves are above the black horizontal line (when the improvement factor is greater than 1). Note that the range in which the FFA outperforms the FFT is larger when searching for lower DM signals. The lines simply connect the data points.

retains all harmonic structure, making it more responsive to pulsars with narrow pulses, the program is not immune to the effect of red noise. As discussed in L15, the fact that the measured sensitivity curves do not match the radiometer curves indicates that these ideal case predictions overestimate the survey sensitivity at low modulation frequencies.

However, in the analysis conducted in L15, the artificial pulsar signals injected in Gaussian noise were also not recovered at the predicted phase-averaged flux densities. This suggests that, perhaps, there are unexpected defects in the FFT search program of the `PRESTO` software. The fact that the improvement from the time-domain search is maximal when the synthetic pulsars have large pulse duty-cycles was also unexpected, as this search type is designed to be particularly efficient at finding signals with large harmonic content. Following this study, the `accelsearch`

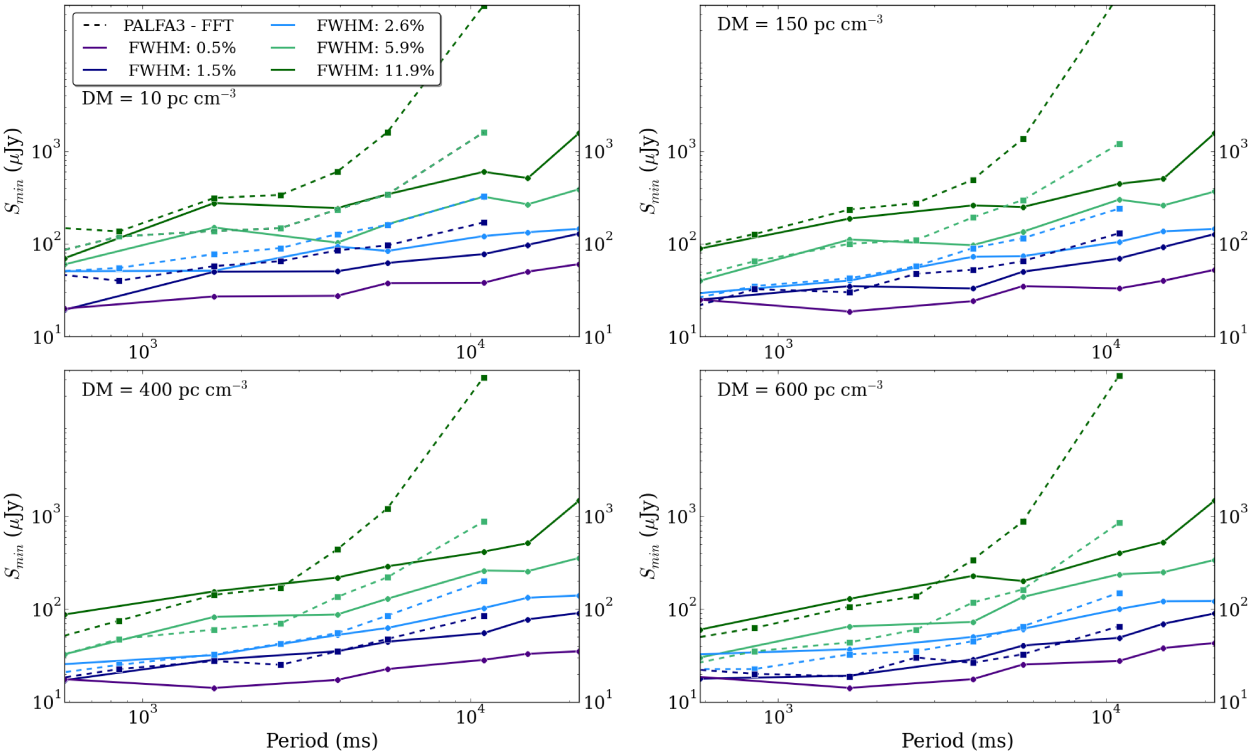


Figure 4.5: Minimum detectable average flux density (S_{min}) curves showing the sensitivity of the FFA in the PALFA survey (solid lines), as measured when using synthetic signals with different pulse widths injected in real survey data. Each panel illustrates the sensitivity curves for different DM values. These signals were injected in only one of the data files, as opposed to the analysis conducted with the FWHM = 2.6 $\%$ signals where the fake pulsars were injected in all 12 data files. One can see that the gain in sensitivity when using the FFA program is more pronounced for pulsars with broad pulses. Note: the lines simply connect the data points.

program was carefully inspected and an error in the code was identified, which was then corrected. The outcome of this issue was that the Fourier-based search looked for the lowest frequencies via the highest harmonics of the fundamental frequency, meaning that the program was insensitive to low frequency signals with little harmonic content. Now that corrections have been implemented in the program, recalculations are being performed to obtain revised sensitivity curves when searching in PALFA data with `accelsearch`. Moreover, recovering those signals is yet more difficult if red noise obscures the fundamental frequency as well as low harmonic frequencies, which may contain a considerable amount of power in the case of pulsars with wide profiles. A Fast-Folding Algorithm has the advantage of including weaker, high harmonic frequencies, and this explains the important gain in sensitivity for broad pulses shown in Figures 4.3 to 4.6.

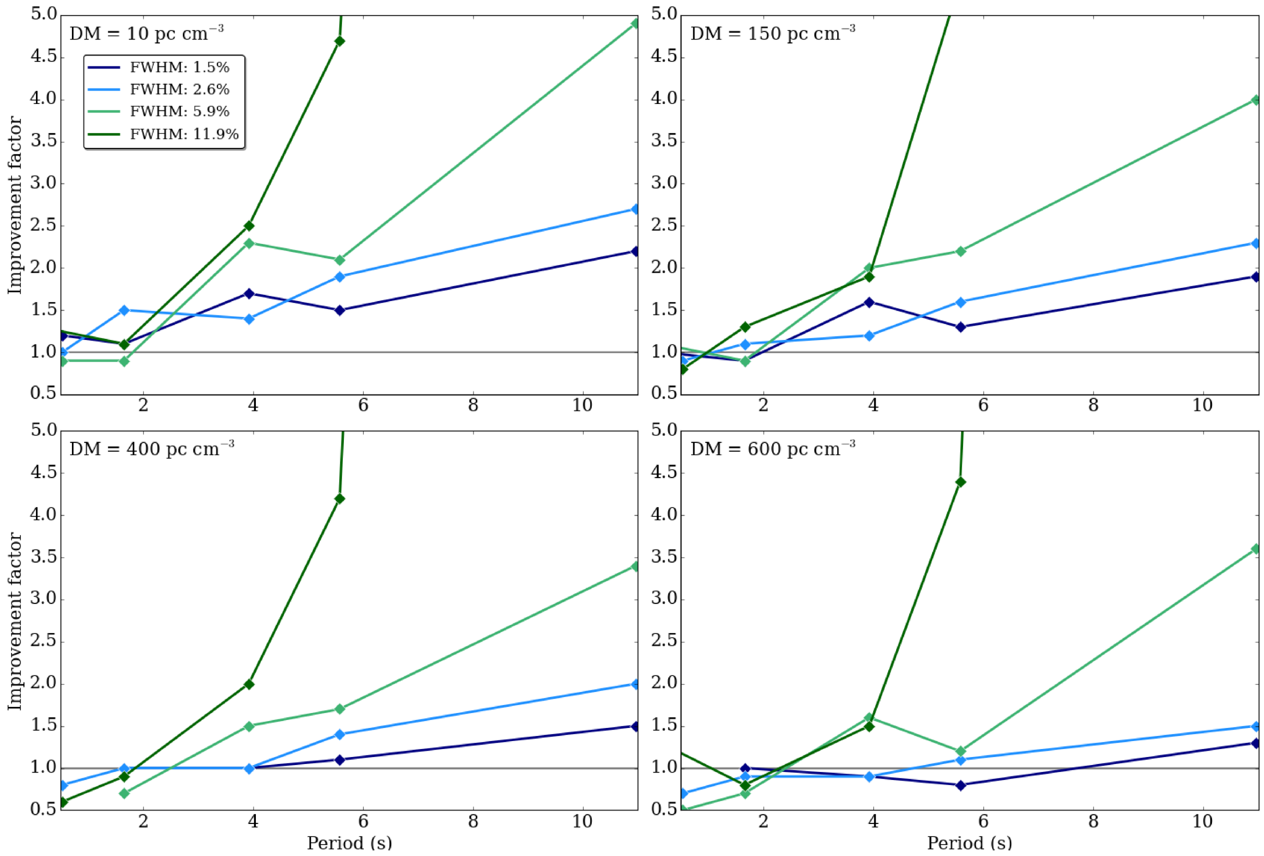


Figure 4.6: Factors of improvement in the sensitivity of the FFA-based search over the FFT-based search as a function of period, for different pulse duty-cycles. Each panel represents a different DM value. As in Figure 4.4, the FFA outperforms the FFT when the curves are above the black horizontal line. Note the high-performance of the FFA compared to the FFT at the lowest DM value (top left panel). The lines simply connect the data points.

To conclude, the sensitivity analysis conducted in this MSc thesis demonstrated the potential of the use of a time-domain search technique in large-scale pulsar surveys for finding long-period pulsars (having either broad or narrow pulses). It would be interesting to reproduce this exercise by constructing pulsar signals with multiple pulse components, such as interpulses, in the same observational data to study the responsiveness of the FFA program to such pulse profiles.

Table 4.4: Improvement factors from the FFA search for different pulse FWHM at a period of 10964.5 sec.

DM (pc cm^{-3})	FWHM (% Phase)	Improvement Factor
	1.5	2.2 ±0.1
10	2.6	2.7 ±0.2
	5.9	4.9 ±0.3
	11.9	62.3 ±3.7
	1.5	1.9 ±0.1
150	2.6	2.3 ±0.1
	5.9	4.0 ±0.2
	11.9	109.5 ±6.6
	1.5	1.5 ±0.1
400	2.6	2.0 ±0.1
	5.9	3.4 ±0.1
	11.9	76.1 ±4.6
	1.5	1.3 ±0.1
600	2.6	1.5 ±0.1
	5.9	3.6 ±0.2
	11.9	82.2 ±4.9

Note the important improvements for signals with FWHM of 11.9 %.

Chapter 5

Conclusions

Finding long-period pulsars can help us understand properties of the true radio pulsar population and could potentially help in constraining models of the mechanism responsible for the emission of the observed radio beams. Such pulsars are also good candidates for neutron star - black hole binary systems, which are yet to be observed. Detecting more radio-loud magnetars, which tend to have long-periods, is another reason why we want to maximize our ability to detect slowly rotating pulsars.

In this MSc thesis, we have presented the implementation of a time-domain search technique for long-period radio pulsars, the Fast-Folding Algorithm (FFA), in the PALFA survey. It is the first time that this algorithm, developed by Staelin in 1969, is used in a large-scale pulsar survey. The extensive computational power required to perform a blind periodicity search in radio pulsar surveys with this algorithm is the main reason why its use has been limited in the past. Now, however, supercomputers have become powerful enough to fulfill this task. We have adapted a FFA program, originally designed to look for transients in Kepler data, so that it can be used to look for pulsar signals in radio data. This adaptation of the program has been recently incorporated in our new version of the pipeline analysis software to search for signals with periods between 500 ms and 30 sec. To date, about 5800 inner Galaxy and 1300 outer Galaxy beams were processed with the new version of the pipeline (which includes beams that were previously analyzed with the older version of the pipeline), and no new pulsars have been discovered via the FFA, yet. This is not surprising, as some of these beams were processed before with the previous version of the PALFA pipeline (i.e., it is likely that if pulsars were present in those observations, they were already discovered from anterior processing) and because there are fewer slowly rotating pulsars in the sample of observed pulsars in general, which is subject to selection effects. Moreover, only a fraction of the produced candidates have been examined so far via the CyberSKA platform, as this process takes a lot of time.

Lazarus *et al.*, (2015) conducted an analysis of the PALFA survey sensitivity by injecting synthetic pulsar signals in real observational data. These signals were then recovered using the Fourier-domain search program included in the pipeline, the `accelsearch` program from the `PRESTO` software package. One of the conclusions from this study is that there is important degradation in the PALFA survey sensitivity in the long-period regime. This reduction in sensitivity was explained by the ruinous effect of red noise, particularly strong at low modulation frequencies. A similar analysis was presented in this MSc thesis, where the same data files containing artificial pulsars were processed by the FFA program to compare the performance of the algorithm against the standard FFT search used in the survey. This work shows an amelioration in the ability of the survey to detect long-period pulsars (periods of $\sim$ 10 sec) by at least a factor of 2, implying that the PALFA survey should discover more long-period radio pulsars in the future. If proven efficient in the survey, the Fast-Folding Algorithm could become a standard pulsar searching technique and could be implemented in other large-scale pulsar surveys, such as the GBNCC Pulsar Survey (see Stovall *et al.*, 2014).

References

1. R. F. Archibald, V. M. Kaspi, S. P. Tendulkar, & P. Scholz, **829**, L21 (Sept. 2016).
2. W. Baade & F. Zwicky, *Proceedings of the National Academy of Science* **20**, 254–259 (May 1934).
3. D. C. Backer, **228**, 42–43 (Oct. 1970).
4. N. D. R. Bhat, J. M. Cordes, F. Camilo, D. J. Nice, & D. R. Lorimer, **605**, 759–783 (Apr. 2004).
5. G. S. Bisnovatyi-Kogan & B. V. Komberg, **18**, 217 (Oct. 1974).
6. R. Browning, D. Ramsden, & P. J. Wright, *Nature Physical Science* **232**, 99–101 (Aug. 1971).
7. J. Chadwick, *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **136**, 692–708, ISSN: 0950-1207 (1932).
8. D. J. Champion *et al.*, *Science* **320**, 1309 (June 2008).
9. W. J. Cocke, M. J. Disney, & D. J. Taylor, **221**, 525–527 (Feb. 1969).
10. J. M. Cordes, **222**, 1006–1011 (June 1978).
11. J. M. Cordes & D. F. Chernoff, **482**, 971–992 (June 1997).
12. S. M. Couch, E. Chatzopoulos, W. D. Arnett, & F. X. Timmes, **808**, L21 (July 2015).
13. H. D. Craft, J. M. Sutton, & J. M. Comella, **219**, 1237–1238 (Sept. 1968).
14. J. S. Deneva *et al.*, **703**, 2259–2274 (Oct. 2009).
15. R. J. Dewey, J. H. Taylor, J. M. Weisberg, & G. H. Stokes, **294**, L25–L29 (July 1985).
16. R. P. Eatough *et al.*, **501**, 391–394 (Sept. 2013).

17. P. C. C. Freire, S. M. Ransom, & Y. Gupta, presented at the 40 Years of Pulsars: Millisecond Pulsars, Magnetars and More, ed. by C. Bassa, Z. Wang, A. Cumming, & V. M. Kaspi, vol. 983, pp. 491–493, DOI: `10.1063/1.2900283`, arXiv: `0711.1883`.
18. F. P. Gavriil *et al.*, *Science* **319**, 1802 (Mar. 2008).
19. R. Giacconi, H. Gursky, F. R. Paolini, & B. B. Rossi, *Physical Review Letters* **9**, 439–443 (Dec. 1962).
20. P. Goldreich & W. H. Julian, **157**, 869 (Aug. 1969).
21. P. A. Henning *et al.*, **139**, 2130–2147 (June 2010).
22. J. W. T. Hessels *et al.*, *Science* **311**, 1901–1904 (Mar. 2006).
23. A. Hewish, S. J. Bell, J. D. H. Pilkington, P. F. Scott, & R. A. Collins, **217**, 709–713 (Feb. 1968).
24. A. W. Hotan, W. van Straten, & R. N. Manchester, **21**, 302–309 (2004).
25. F. Hoyle, J. V. Narlikar, & J. A. Wheeler, **203**, 914–916 (Aug. 1964).
26. R. A. Hulse & J. H. Taylor, **195**, L51–L53 (Jan. 1975).
27. C. Kiddle *et al.*, presented at the Astronomical Data Analysis Software and Systems XX, ed. by I. N. Evans, A. Accomazzi, D. J. Mink, & A. H. Rots, vol. 442, p. 669.
28. M. M. Komesaroff, **225**, 612–614 (Feb. 1970).
29. V. I. Kondratiev *et al.*, **702**, 692–706 (Sept. 2009).
30. M. Kramer, **509**, 856–860 (Dec. 1998).
31. L. D. Landau, *Phys. Zs. Sowjet., vol.1, p.285, 1932 (English and German)* **1**, 285 (Dec. 1932).
32. J. M. Lattimer, presented at the American Institute of Physics Conference Series, ed. by S. Kubono *et al.*, vol. 1484, pp. 319–326, DOI: `10.1063/1.4763414`.
33. M. A. Livingstone *et al.*, **308**, 317–323 (Apr. 2007).
34. D. R. Lorimer & M. Kramer, *Handbook of Pulsar Astronomy*.
35. B. Y. Losovsky & D. V. Dumsky, *Astronomy Reports* **58**, 537–544 (Aug. 2014).
36. K. Mori *et al.*, **770**, L23 (June 2013).
37. S. A. Olausen & V. M. Kaspi, **212**, 6 (May 2014).
38. J. R. Oppenheimer & G. M. Volkoff, *Physical Review* **55**, 374–381 (Feb. 1939).

39. F. Pacini, **216**, 567–568 (Nov. 1967).

40. W. Pauli, *Zeitschrift für Physik A Hadrons and Nuclei* **36**, 336–363 (May 1926).

41. D. J. Price & S. Rosswog, *Science* **312**, 719–722 (May 2006).

42. V. Radhakrishnan & C. S. Shukre, presented at the Supernovae, their Progenitors and Remnants, ed. by G. Srinivasan & V. Radhakrishnan, p. 155.

43. S. M. Ransom *et al.*, **505**, 520–524 (Jan. 2014).

44. J. P. Ridley *et al.*, **433**, 138–146 (July 2013).

45. M. A. Ruderman & P. G. Sutherland, **196**, 51–72 (Feb. 1975).

46. P. A. G. Scheuer, **218**, 920–922 (June 1968).

47. P. Scholz *et al.*, **833**, 177 (Dec. 2016).

48. M. Schönberg & S. Chandrasekhar, **96**, 161 (Sept. 1942).

49. L. G. Spitler *et al.*, **531**, 202–205 (Mar. 2016).

50. D. H. Staelin, *IEEE Proceedings* **57**, 724–725 (1969).

51. I. H. Stairs, A. G. Lyne, & S. L. Shemar, **406**, 484–486 (Aug. 2000).

52. K. Stovall *et al.*, **791**, 67 (Aug. 2014).

53. T. M. Tauris, N. Langer, & M. Kramer, **416**, 2130–2142 (Sept. 2011).

54. D. J. Thompson, C. E. Fichtel, D. A. Kniffen, & H. B. Ogelman, **200**, L79–L82 (Sept. 1975).

55. R. C. Tolman, *Physical Review* **55**, 364–373 (Feb. 1939).

56. J. van Leeuwen *et al.*, **798**, 118 (Jan. 2015).

57. P. T. Wallace *et al.*, **266**, 692–694 (Apr. 1977).

58. N. Wang, R. N. Manchester, & S. Johnston, **377**, 1383–1392 (May 2007).

59. M. D. Young, R. N. Manchester, & S. Johnston, **400**, 848–849 (Aug. 1999).

60. B. Zhang, A. K. Harding, & A. G. Muslimov, presented at the American Astronomical Society Meeting Abstracts #195, vol. 32, p. 880.