

UNIVERSITY OF SOUTHAMPTON
FACULTY OF SOCIAL, HUMAN AND MATHEMATICAL SCIENCES
GRAVITY GROUP

**Timing variations in neutron stars:
models, inference, and their implications
for gravitational waves**

Author:
Gregory ASHTON

Supervisors:
Dr. David Ian JONES
Dr. Reinhard PRIX

Thesis for the degree of Doctor of Philosophy

July 21, 2016

UNIVERSITY OF SOUTHAMPTON

ABSTRACT

FACULTY OF SOCIAL, HUMAN AND MATHEMATICAL SCIENCES
Gravity Group

Doctor of Philosophy

TIMING VARIATIONS IN NEUTRON STARS: MODELS, INFERENCE, AND
THEIR IMPLICATIONS FOR GRAVITATIONAL WAVES

by Gregory Ashton

Timing variations in pulsars, low frequency ubiquitous structure known as timing noise and sudden increases in the rotational frequency which we call glitches, provide a means to study neutron stars. Since the first observations, many models have been proposed, yet no definitive explanation has arisen.

In this thesis, we aim to improve this situation by developing models of timing noise. We focus chiefly on precession models which explain periodic modulation seen in radio pulsar data. Developing models and testing them provides an opportunity to infer the elemental properties of neutron stars: evidence for long period precession has implications for the superfluid component predicted by models used to explain glitches. However, often more than one model can qualitatively explain the data, therefore we need a method to decide which model best fits the data. This is precisely the case for PSR B1828-11 which has been used as evidence for both precession and so-called magnetospheric switching. We address this confusion by applying the tools of probability theory to develop a Bayesian model comparison and find that the evidence is in favour of precession.

In the second part of this thesis, we will discuss the implications of timing variations for the detection of continuous gravitational waves from neutron stars. To search for these signals, matched filtering methods are used which require a template, a guess for what the signal ‘looks like’. Timing variations, as seen in the electromagnetic signal, may also exist in the gravitational wave signal. If detected, these could provide an invaluable source of information about neutron stars. However, if not included in the template, they may mean that the gravitational wave signal is not detected in the first place. We investigate this issue for both timing noise and glitches, using electromagnetic observations to predict for what types of gravitational wave searches this may be an issue. We find that while timing noise is unlikely to be an issue for current gravitational wave searches, glitches may cause a significant problem in all-sky searches for gravitational waves from neutron stars.

*One man's 'noise' is
another man's 'signal'.*

EDWIN THOMPSON JAYNES

Contents

Declaration of Authorship	xi
Acknowledgements	xiii
1 Introduction	1
1.1 Observation of pulsars and their identification with neutron stars	2
1.2 Pulsar timing	3
1.3 Categorising neutron stars	5
1.4 The physics of rotation powered pulsars	7
1.5 Radio pulsar population statistics	10
1.6 Neutron stars and gravitational waves	11
1.7 Plan of the thesis	14
2 Timing variations	15
2.1 Glitches	15
2.2 Timing noise: observations	18
2.3 Timing noise: interpretations	20
2.3.1 Random walk models	20
2.3.2 Free precession	23
2.3.3 Magnetospheric switching	25
Appendix 2.A Toy model of stochastic resonance: particle in a potential . . .	33
3 Action of the electromagnetic torque on a precessing neutron star	35
3.1 Defining the model	36
3.2 Spherical star	39
3.3 Biaxial neutron star with no anomalous torque	39
3.3.1 Neutron star A in region $\tau_S \gg \tau_P$	41
3.3.2 Neutron star B in region $\tau_S \sim \tau_P$	46
3.3.3 Neutron star C in region $\tau_S \ll \tau_P$	46
3.4 Biaxial neutron star including the anomalous torque	47
3.4.1 The effective body frame	47
3.4.2 Phase space	50
3.4.3 Neutron star A in region $\tau_S > \tau_A > \tau_P$	51
3.4.4 Neutron star B in region $\tau_S > \tau_P > \tau_A$	53
3.4.5 Neutron star C in region $\tau_P > \tau_S > \tau_A$	55
3.5 Conclusion	55
Appendix 3.A Considerations for the timescales	56

4 Modelling observations of precessing pulsars	59
4.1 Rotating into the inertial frame	60
4.1.1 Euler rotation matrices	60
4.1.2 Evolution of the Euler angles	61
4.1.3 Initial conditions	62
4.2 Evolving the Euler angles	63
4.2.1 Torque free biaxial body	64
4.2.2 Torqued biaxial body	64
4.3 Precessing pulsars	67
4.3.1 The reference plane	67
4.3.2 Dynamics of the magnetic dipole	68
4.3.3 Understanding the dynamics of the magnetic dipole	70
4.3.4 The effective wobble angle	72
4.4 Observable features: the phase residual	73
4.4.1 Effect of precession on the phase residual	74
4.4.2 Effect of torqued precession on the phase residual: electromagnetic amplification	76
4.5 Observable features: the spin-down rate	77
4.5.1 Derivation of the precession spin-down rate	78
4.5.2 Simulations of the precession spin-down rate	80
4.6 Observable features: the pulse profile	82
4.6.1 Variations in the pulse intensity	83
4.6.2 Variations in the beam-width	84
4.7 Application: switching and precession	87
4.7.1 Switching in the spin-down torque only	90
4.7.2 Switching in the spin-down and anomalous torque	91
4.8 Conclusions	94
5 Comparing models of the periodic variations in spin-down and beam-width for PSR B1828-11	95
5.1 Introduction	95
5.2 Bayesian Methodology	98
5.2.1 The odds-ratio and posterior probabilities	98
5.2.2 Signals in noise	99
5.2.3 Choosing prior distributions	100
5.3 Defining and fitting the models	101
5.3.1 Noise-only model	102
5.3.2 Switching model	103
5.3.3 Precession model	111
5.4 Estimating the odds-ratio	125
5.4.1 Thermodynamic integration	125
5.4.2 Results	126
5.4.3 Effect of the choice of prior	127
5.5 Discussion	129
Appendix 5.A Procedure for MCMC parameter estimation	130
Appendix 5.B Implications for the unobserved beam	131

6 Continuous gravitational waves: calculating the mismatch	133
6.1 Introduction	133
6.2 Introduction to the mismatch	134
6.2.1 Defining the mismatch	135
6.2.2 Interpreting the mismatch	135
6.2.3 Taylor expansion signals and templates	136
6.2.4 The metric-mismatch approximation for fully-coherent searches . .	137
6.2.5 The mismatch for semi-coherent searches	137
6.3 Exact mismatch from irregularities in the phase	139
6.3.1 Two subdomains with a phase discontinuity	139
6.3.2 N subdomains with phase discontinuities	140
6.3.3 Oscillating phase deviations	142
6.4 Generalising the metric-mismatch approximation to arbitrary signals . .	143
6.4.1 Piecewise Taylor expansion	143
6.4.2 The generalised metric-mismatch approximation for fully-coherent searches	144
6.4.3 Explicit calculation of the metric	146
7 Glitches in continuous gravitational waves	149
7.1 Introduction	149
7.2 Continuous gravitational-wave searches	151
7.3 Statistical properties of the observed glitch database	152
7.3.1 Glitch magnitudes	153
7.3.2 Overview of the population of glitches	154
7.3.3 Extrapolating: glitch magnitudes	156
7.3.4 Extrapolating: average glitch rate	158
7.4 Calculating the mismatch due to a single glitch	161
7.4.1 A single glitch in a fully-coherent search	163
7.4.2 A single glitch in a semi-coherent search	165
7.5 Predicting the mismatch and rate of glitches in gravitational wave searches	168
7.5.1 Fully-coherent searches	169
7.5.2 Semi-coherent searches	170
7.5.3 The follow-up stage	172
7.5.4 Including the recovery from glitches	173
7.5.5 Application to past and future searches	174
7.6 Conclusions	175
Appendix 7.A Bayesian model comparison: test of mixture models	176
Appendix 7.B Linear regression in log-space	179
Appendix 7.C Understanding the uncertainty in the predictions	179
8 Timing noise in continuous gravitational waves: a numerical study	183
8.1 Introduction	183
8.2 Timing noise as described by the Crab ephemeris	186
8.3 Method	187
8.4 Results	189
8.4.1 The effect of timing noise on narrow-band searches	189
8.4.2 Results relevant to recent narrow-band searches	190

8.4.3	Minimum mismatch as a function of the observation epoch	191
8.4.4	Averaged minimum mismatch as a function of the observation duration	193
8.5	Conclusions	195
9	Timing noise in continuous gravitational waves: random walk models	197
9.1	Defining a random walk	198
9.1.1	Initial definitions	198
9.1.2	Writing the parameter offsets in terms of normal distributions . .	200
9.2	Random walk models: a simple treatment	201
9.2.1	Taking the expectation	202
9.2.2	Verifying the results	202
9.3	Random walk models: minimising the mismatch	203
9.3.1	Random walk in the phase	204
9.3.2	Random walk in the frequency	206
9.3.3	Verifying the results	207
9.4	Understanding the random walk model	207
9.5	Application to the Crab pulsar	210
9.5.1	Distribution of jumps in the Crab ephemeris	211
9.5.2	Predicting the mismatch in the Crab	213
9.6	Conclusion	214
Appendix 9.A	Summation identities	215
Appendix 9.B	Least-squares minimisation of a random walk	216
9.B.1	Least squares fitting of a polynomial	216
9.B.2	Least squares fitting a polynomial to a random walk	217
9.B.3	Zeroth order fitting	217
9.B.4	First order fitting	219
9.B.5	Second order fitting	221
9.B.6	Conclusions	221
10	Conclusion and outlook	223

Declaration of Authorship

I, Gregory Ashton, declare that the thesis entitled *Timing variations in neutron stars: models, inference, and their implications for gravitational waves* and the work presented in the thesis are both my own, and have been generated by me as the result of my own original research. I confirm that:

- this work was done wholly or mainly while in candidature for a research degree at this University;
- where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
- where I have consulted the published work of others, this is always clearly attributed;
- where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
- I have acknowledged all main sources of help;
- where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
- parts of this work have been published as: [24] and [25].

Signed:.....

Date:.....

Acknowledgements

I would like to thank Ian Jones, my supervisor in Southampton, for his tireless efforts to guide my research in a meaningful direction, his clear and honest evaluation, but most importantly for being a friendly and enthusiastic mentor. Reinhard Prix, my supervisor in Hannover, also deserves a hearty thanks for his enduring commitment, many useful insights, and for introducing me to Bayesian probability theory, without which much of this thesis would not exist. It has been a pleasure to work with both Ian and Reinhard. Thanks also to my advisor, Nils Andersson, for his useful input early on and his continued humour throughout such as his brilliant suggestion to title the work ‘A waste of timing noise’. To my examiners, Graham Woan and Wynn Ho, I extend my thanks for taking the time to provide detailed and insightful comments which I thoroughly appreciate.

For the work presented in Chapter 5, I am grateful to Will Farr, Danai Antonopoulou, and Ben Stappers for valuable discussions and comments, Dan Foreman-Mackay for the software used in generating posterior probability distributions [61], and Andrew Lyne [111] for generously sharing the data for PSR B1828-11. For the work presented in Chapter 7, I kindly thank Christobal Espinoza for maintaining the glitch catalogue and for helping to extract the required data. I would also like to express my gratitude to Matthew Pitkin whose advice on MCMC simulation software was of great help.

All of my friends in the Maths departments deserve my praise both for their moral support and making Southampton an enjoyable place to live and work. To my final year companions, Vanessa Gruber, Marta Colleoni, Liana Kontogeorgaki, and Yafet Sanchez, I have thoroughly enjoyed the adventure we had together and I wish you all the best of luck in your future pursuits. I particularly would like to acknowledge Yafet for his willingness to question everything and reminding me that while life may be difficult, consider the alternative. Matthew Cobain, Marc Scott and Andrew Meadowcroft, I thank for their kind help in editing this thesis, many hours of badminton, many enjoyable discussions, and many gin & tonics.

My heartfelt thanks to Emily Lawrence for her solid support, wonderful cooking, and for keeping me sane; to Christine Ashton for her everlasting enthusiasm and words of wisdom; and finally to the man who inspired me to such heights, Mark Ashton.

Chapter 1

Introduction

Neutron stars were first postulated by Landau as ‘dense stars which look like giant atomic nuclei’ [171], even before the discovery of the neutron by Chadwick [38]. However, it was Baade and Zwicky [26] who made the explicit prediction of a neutron star whilst trying to explain the energy released in observed supernova.

In a main-sequence star, the nuclear fusion of hydrogen atoms into helium provides outward pressure balancing the star in an equilibrium configuration with the inward pressure of the star’s self-gravity. Eventually the star depletes its reserves of hydrogen and can no longer maintain equilibrium. If the star has an initial mass greater than $\sim 8M_{\odot}$, then it may undergo a *core-collapse supernova* during which some of the mass is ejected, but the rest falls in creating a new compact object. In this object, temperatures and pressure rapidly soar and the electrons and protons undergo inverse beta decay combining to form neutrons and neutrinos:

$$e^- + p \rightarrow n + \nu. \quad (1.1)$$

If the compact object’s mass is less than the maximum mass of a stable neutron star [130], then, once the pressure reaches nuclear densities of $\sim 2.3 \times 10^{14} \text{ g/cm}^3$, neutron degeneracy pressure can halt the collapse in a new compact stable equilibrium configuration which we call a neutron star. The exact value of the maximum mass depends on the equation of state of matter at high densities, but typical values range from 1.5 to $3 M_{\odot}$ [32]. For remnants with larger masses, this is not possible and the object will collapse to form a black-hole; the detail of exactly what the critical mass a neutron star can sustain is sensitive to the equation of state of matter under these conditions.

Our knowledge of neutron stars is founded on observations made by electromagnetic astronomy. This has revealed a wealth of different neutron stars and their phenomena which we will introduce in the following sections. Many of these observed phenomena have well defined models which allow us to infer properties of neutron stars and their environments. However, our knowledge of neutron stars is far from complete: current interpretations of the observations can be contradictory or have features not explained by any known physical models. Improvements in electromagnetic astronomy will bring

to light a greater number of new neutron stars and improve the resolution of those currently observed; it is hoped that this will help us to better understand them.

There are two other methods we can utilise to learn more about neutron stars: improved modelling of current observations and by observing them from their gravitational wave emissions. In this thesis, we will study how one of the observed phenomena, so-called ‘timing variations’, can help us to learn more from current observations and also test whether it may hinder the current search for gravitational waves from neutron stars.

In this introduction, we will acquaint the reader with the current observations of neutron stars and introduce some basic physics.

1.1 Observation of pulsars and their identification with neutron stars

After the conception of neutron stars as stable compact objects there was thought to be little chance of observing them. They are many orders of magnitude smaller than other celestial objects and soon after their formation in rare supernovae events (van den Bergh [160] predicts just 3 supernovae per century in the Milky Way Galaxy) they are rapidly cooled by the emitted neutrinos making their thermal emission difficult to detect.

In 1968 a bright periodic electromagnetic (EM) signal, now known as PSR B1919+21, was identified by Hewish et al. [78] during a high time-resolution survey for interplanetary scintillation. The source was measured with a radio frequency of 81.5 MHz and pulsed with a period of ~ 1.377 s; this led to the name *pulsars* to refer to such sources. Following this, several other similar objects were discovered. A unifying feature of all pulsars is the clock-like stability of the pulsations - something not rivalled by any other astrophysical phenomenon. This stability indicates that the source must be a collimated beam fixed to a rotating body such that, as it rotates, the beam sweeps out like a lighthouse; in this way the pulsation period is exactly the rotation period of the body. Alternative models such as emission due to accretion from a binary companion could never reach the stability’s seen in pulsars.

For any rotating body to remain gravitationally bound, its rotation frequency is constrained by the requirement that the centrifugal acceleration at the equator be less than the gravitational acceleration. The rotation frequency of the observed pulsars ruled out all known astrophysical bodies except the two most compact objects, neutron stars and black holes, since all other bodies would not be gravitationally bound at these frequencies. Isolated black holes are unable to support an electromagnetic emission mechanism. This left only neutron stars as candidates.

The explicit identification of pulsars with neutron stars came from Gold [68], however, Pacini [131] had already predicted the basic ingredients of the model before the observation. They suggested that a rapidly rotating neutron star with a strong dipolar magnetic field would stream radiation out along the magnetic axis. If this axis is misaligned from the rotation axis, then the beams are swept out like a lighthouse. Beams passing over

the earth are observed as periodic pulses at the rotation frequency of the star. Such a model predicts that the electromagnetic radiation should exert a torque and slow-down the rotation. This slowdown was subsequently measured in the Crab pulsar, which was discovered at the centre of the Crab nebulae, a supernova remnant, agreeing with the prediction of Baade and Zwicky [26].

These early detections gave birth to a new field of astronomy: pulsar astronomy. Since then, researchers in the field have detected over 2000 radio pulsars, measured thermal emission from a handful of nearby neutron stars with typical temperatures of 10^5K to 10^6K [133], found neutron stars in accreting binary systems with companions, and identified many other ways to observe neutron stars. We will discuss some of these which are relevant to this thesis in Section 1.3, but first in Section 1.2 we describe the techniques used by the field to identify and ‘time’ radio pulsars.

1.2 Pulsar timing

Pulsars can be observed by measuring the variation in amplitude of the radio waves from a particular sky location. A single observation consists of measuring the amplitude over a time period of approximately 30 minutes or so, which, for pulsars with spin period $\sim 1\text{ s}$, means recording up to several thousand individual pulsations.

The shape of individual pulses can vary substantially during a single observation; to demonstrate this, in Figure 1.1, successive pulses from PSR B1919+21, the first discovered pulsar, are vertically stacked and aligned. Each pulsation lasts for a small fraction of the pulse period. As an example, the pulses in PSR B1919+21, shown in Figure 1.1, have typical widths of 0.031 s, but the pulse period is 1.337 s; note that the stacked plot is truncated to show only the pulsation itself. Later, in Figure 1.6, we will show this to be a common characteristic for the normal radio pulsar population by looking at the duty-cycle, the ratio of the pulse duration to the pulse period.

In order to understand the gross features of a pulsar, astronomers average over the hundreds to thousands of pulses observed during a single observation to create a single integrated pulse profile. This is done by sampling the radio signal at fixed time intervals then ‘folding’ all the samples at the pulse period (for a complete review see Chapter 15 of Lyne and Graham-Smith [110]). In contrast to the individual pulses which, as shown in Figure 1.1, can be highly variable, the integrated pulse profile is highly stable between independent observations over timescales of years.

The integrated pulse profile not only gives a stable picture of what the pulsations look like on average, but it also provides a highly accurate measurement of the time of arrival (TOA) of a single pulse during the observation. It is this TOA which can be used to ‘time’ a pulsar. To do this, regular observations of a pulsar must be made every few months or so, with each observation resulting in a precise TOA measurement. Having obtained a series of TOAs, pulsar astronomers generate a *timing model* which attempts to exactly count each and every pulse. Between any two observations there may be several million pulses so the timing model needs to account for any mechanisms which may produce

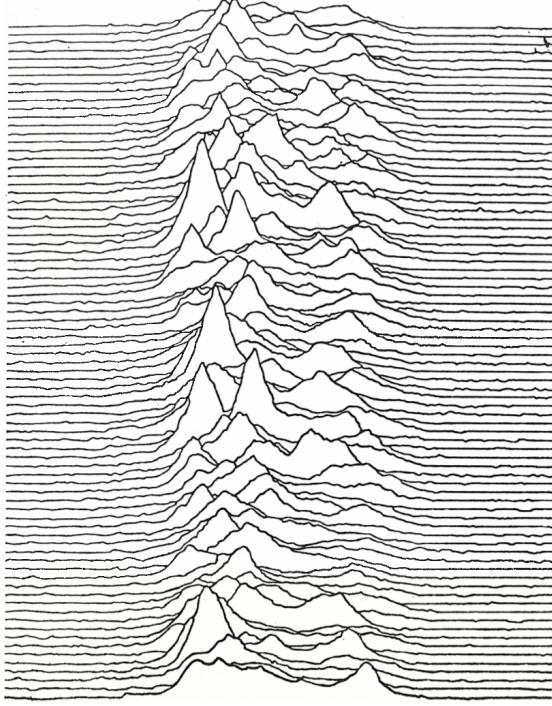


Figure 1.1: The radio amplitude of successive pulses from PSR B1919+21 stacked vertically, figure reproduced from Mitton [126], originally produced by Craft [50].

variations in the TOAs. The process is standardised by the software package **TEMPO2** developed by Hobbs et al. [80], of which we will now describe the essential features.

The TOA of a pulse at the detector on Earth depends on many factors, such as: the time at which the beam was directed by the source towards the Earth; the relative motions of the source and detector; and any mechanisms effecting the signal during its transit. In this thesis, we will be concerned only with the time at which pulses are generated (when the source beams towards the earth) which is governed by the *timing properties* of the star itself. As described by Edwards et al. [57] these can be modelled by a Taylor expansion in the phase at time t , given by

$$\phi(t) = \phi_0 + 2\pi \sum_{n \geq 1} \frac{\nu^{(n-1)}}{n!} (t - t_{\text{ref}})^n, \quad (1.2)$$

where $\{\nu^{(n)}\}$ is the n^{th} time derivative of the spin frequency of the rotating body, ϕ_0 is the initial phase, and t_{ref} is an arbitrary reference time. This expansion is usually truncated at $n = 3$, the second order spin-down rate $\ddot{\nu}$. The timing model then includes corrections to this to model the relative motion of the source and detector, intergalactic transit, and other effects; these are described in full in Edwards et al. [57].

Between any two TOAs, if the timing model is correct, an integer number of rotations must have occurred; this allows the use of the deviation of $\phi(t_{\text{TOA}}^j)$ from an integer as a test statistic. The timing model minimises the root-mean-square (RMS) of these deviations with respect to the timing model parameters, for example the frequency and

frequency derivatives. The output of applying a particular timing model (choice of corrections) to a set of data is then the best-fit of these parameters and an estimate of their associated errors. The corrections applied in a timing model provide a method to investigate pulsar physics: for example in some pulsars an orbital correction must be applied which models the periodic motion of the star due to an orbital companion. Using this technique, Wolszczan and Frail [170] discovered the first *exoplanets* orbiting the pulsar PSR B1257+12.

A minimisation of the timing model parameters will converge regardless of whether or not the model itself is appropriate. To qualitatively check if the fitted model described the data, pulsar astronomers refer to the *timing residual*, which is the difference between the TOA, as given by the timing model, and the actual TOA. The timing residual provides a mechanism to evaluate the timing model: for example a periodic variation in the timing residual with period 365 days may indicate the correction of the Earth's orbit about the Sun may be incorrect. If the timing model is correct, the residual data points should be Gaussianly distributed around zero. A timing model is described as *phase-connected* if it is accurate enough to track the pulsar to within a single rotation. For most pulsars this is the case and a single set of coefficients can track the spin-down over periods greater than a year.

However, for all pulsars the timing residual contains ‘structure’ known as *timing variations* which cannot be associated with any known correction. These variations are the focus of this work and we will describe the details further in Section 2. In the next section, we will describe the variety of known pulsars which have been timed using this method.

1.3 Categorising neutron stars

The timing properties, and other features measured by the timing model, for over 2000 pulsars observed can be accessed via the Australia Telescope National Facility (ATNF) pulsar catalogue [117]. We can categorise the population by their measured values of period P and the period derivative \dot{P} . This is done by plotting them in a so-called $P - \dot{P}$ diagram, as shown in Figure 1.2. The various categories to which each pulsar can be assigned have been marked in this plot and we now discuss their features.

The majority of pulsars (referred to as the ‘normal’ pulsars) are found *isolated* without a binary companion and have typical periods of $P = 10^{-2} - 10^1$ s. These can be described as *rotation powered* pulsars, since the electromagnetic (EM) radiation is powered by the loss of rotational energy. As described later in Section 1.4, estimates can be made of their characteristic age, τ_{age} , and surface magnetic field strength, B_0 , based on a dipole spin-down model. Constant lines of these quantities are plotted in Figure 1.2. Of the normal pulsars, we can choose to identify the young pulsars as those for which $\tau_{\text{age}} < 10^5$ yrs. Some of these, such as the Crab and Vela pulsars, can be directly associated with their supernova remnant from which they were formed [95].

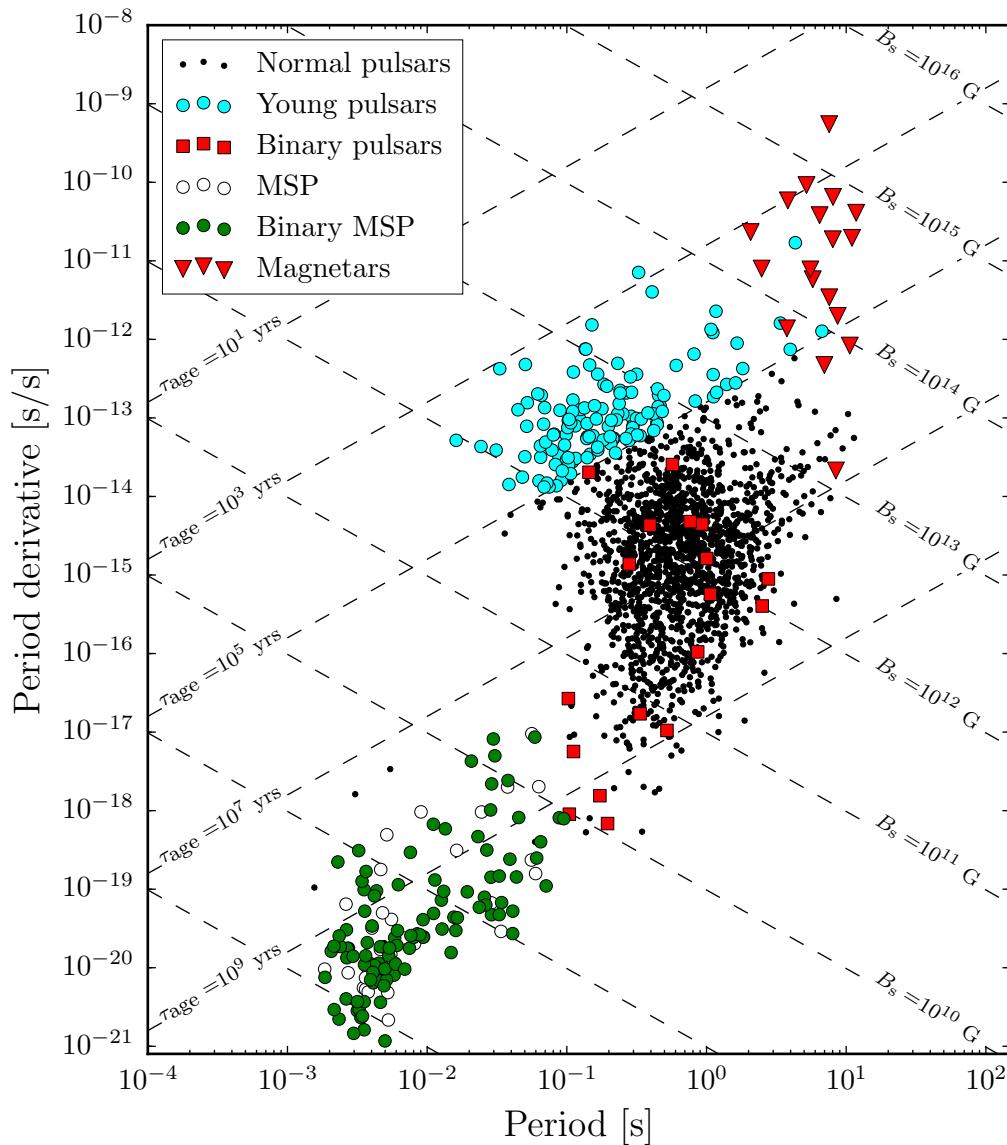


Figure 1.2: Period - Period derivative diagram using data taken from the ATNF pulsar catalogue [117]. Dashed lines show inferred characteristic ages and surface magnetic fields as given by Eqn. (1.10) and Eqn. (1.13) respectively.

A second smaller population of isolated rotation powered pulsars exists with $P < 10^{-1}$ s. These are known as the *millisecond pulsars* (MSPs). This special class of pulsars are believed to start their life as normal pulsars, but are then spun-up through accretion from a normal star. In support of this hypothesis, the majority of MSPs in Figure 1.2 have a binary companion [169]. Additionally, we see so-called low-mass X-ray binary systems (LMXBs) which are systems where a neutron star in a binary accretes matter from its companion; the infalling matter releases gravitational potential energy in the form of X-rays (see for example pg. 73 of Lewin et al. [105]). It is thought that these LMXBs are the progenitors of the MSPs; recent results of ‘transitional systems’ (see for example [20]) which switch between the two, appear to confirm this.

We include one final class of neutron stars, *magnetars*, thought to have large magnetic fields of $B \gtrsim 5 \times 10^{13}$ G. These are in fact observed from two channels which we will now describe. Some pulsars are observed to emit X-ray radiation; usually this is powered by the accretion of matter from a binary companion, but this mechanism does not apply to isolated stars. Subsequently, isolated stars observed in the X-ray band were named *anomalous X-ray pulsars* (AXPs). It was shown by Duncan and Thompson [56] that AXPs are magnetars where the emission is powered by the decay of the strong magnetic field. At the same time, astronomers found a class of objects emitting irregular bursts of γ -rays or X-rays which they named the *soft γ -ray repeaters* (SGRs). As discussed in Kouveliotou et al. [96] these are now understood to be magnetars which undergo rearrangement of their magnetic fields. The two individual observations were unified by observation of X-ray bursts from AXPs by Gavriil et al. [63]. In Figure 1.2 we label observations from both these sources as magnetars.

1.4 The physics of rotation powered pulsars

For rotation powered pulsars, the normal population in Figure 1.2, we can infer a substantial amount about their physics by applying a simple model to their observed timing properties. In this section, we will introduce such a model and acquaint the reader with methods to infer the spin-down ages and magnetic fields, important quantities in understanding neutron stars.

Let us model the star as described by Pacini [131] and Gold [68] and illustrated in Figure 1.3: a rapidly rotating body with a magnetic dipole fixed in the crust at an angle α to the rotation axis.

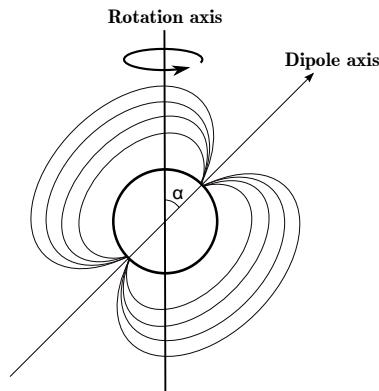


Figure 1.3: An illustration of the dipole spin-down model. The dipole and some of the closed field lines are fixed at an angle α to the rotation axis. As the body rotates, radiation is emitted along both ends of the dipole axis producing a torque on the body.

From §67 of Landau and Lifshitz [101], the total radiation from a dipole rotated at angular frequency Ω can be shown to be given by

$$\Upsilon = \frac{2}{3} \frac{\Omega^4}{c^3} d_0^2, \quad (1.3)$$

where d_0 is the projection of the dipole moment on the plane perpendicular to the axis of rotation [131]. We note that we could alternatively parameterise with the pulse period $P = \frac{2\pi}{\Omega}$. Following the arguments in §10.5 of Shapiro and Teukolsky [154], the magnitude of the magnetic dipole moment for a star with radius R and surface magnetic field strength B_0 is $B_0 R^3 / 2$. Including the projection onto the plane perpendicular to the rotation axis, the total radiation is then

$$\Upsilon = \frac{1}{6} \frac{\Omega^4}{c^3} B_0^2 R^6 \sin^2 \alpha. \quad (1.4)$$

The rotational energy of a body spinning at Ω with a moment of inertia I_0 is given by

$$E = \frac{1}{2} I_0 \Omega^2. \quad (1.5)$$

Differentiating this expression with respect to time gives the loss of rotational energy, $\dot{E} = I_0 \Omega \dot{\Omega}$ where $\dot{\Omega}$ is the angular *spin-down rate* which can be related to the changing pulse period by $\dot{P} = -2\pi \frac{\dot{\Omega}}{\Omega^2}$. Assuming that all the energy is lost to the rotation of the dipole, hence the name rotation powered pulsars, we can equate $\dot{E} = -\Upsilon$. We then rearrange to give a power-law relation between the spin-down rate and the spin-frequency:

$$\dot{\Omega} = -\frac{B_0^2 R^6 \sin^2 \alpha}{6c^3 I_0} \Omega^3. \quad (1.6)$$

This power-law dependence is a model specific version of a more general phenomenological power-law braking model

$$\dot{\Omega} = -k \Omega^n. \quad (1.7)$$

Generalising in this way suggests a powerful method to determine the type of braking for a given pulsar. Specifically, differentiating Eqn. (1.7) and rearranging it can be shown that

$$n = \frac{\ddot{\Omega} \Omega}{\dot{\Omega}^2}. \quad (1.8)$$

Therefore, if $\ddot{\Omega}$ can be measured, then n can be determined, and hence used to infer the type of braking. For example, measuring $n = 3$ would indicate the pulsar braking is dominated by losses due to the magnetic dipole, in contrast, it can be shown that gravitational wave braking would produce $n = 5$ [154, pg.284]. Unfortunately, in reality, pulsars do not constrain this value. Work by Biryukov et al. [29] found (see Figure 1.4) that younger pulsars tend to have braking indices of the correct order of magnitude. However, beyond $\tau_{ch} \approx 10^5$ years the absolute value of the braking index rapidly grows, reaching values as large as 10^6 for the oldest pulsar. In addition, an almost equal number of pulsars have positive and negative values of the braking index.

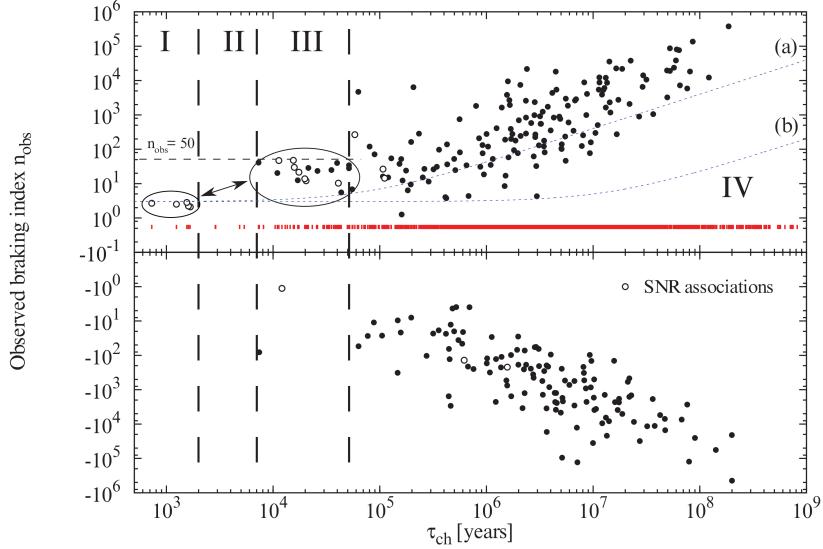


Figure 1.4: The inferred observed braking index n_{obs} against the characteristic age $\tau_{\text{ch}} = \tau_{\text{age}}$ values for 1337 ordinary rotation powered pulsar, figure reproduced from Biryukov et al. [29].

To infer the age of the pulsar, Eqn. (1.7) can be integrated between the initial values ($t = 0, \Omega = \Omega_i$) and the observed value (Ω_o) to give

$$t = \frac{1}{(1-n)} \frac{\Omega_o}{\dot{\Omega}_o} \left(1 - \frac{\Omega_o^{n-1}}{\Omega_i^{n-1}} \right). \quad (1.9)$$

Typically, we make the assumption that all pulsars, regardless of their measured braking index, are dominated by EM braking such that $n = 3$. Then additionally assuming that $\Omega_i \gg \Omega_o$ we can approximate to a characteristic age

$$\tau_{\text{age}} = -\frac{\Omega_o}{2\dot{\Omega}_o} = \frac{P}{2\dot{P}}. \quad (1.10)$$

To infer the approximate surface magnetic field strength, we first note that in the EM dipole braking model:

$$k = \frac{B_0^2 R^6 \sin^2 \alpha}{6c^3 I_0}. \quad (1.11)$$

Then rearranging Eqn. (1.6) we can estimate the surface magnetic field strength at the poles by

$$B_0 = \left(\frac{6c^3 I_0}{R^6 \sin^2 \alpha} \right)^{\frac{1}{2}} \left(\frac{-\dot{\Omega}}{\Omega^3} \right)^{\frac{1}{2}} = \frac{1}{2\pi} \left(\frac{6c^3 I_0}{R^6 \sin^2 \alpha} \right)^{\frac{1}{2}} \sqrt{P\dot{P}} \quad (1.12)$$

In general we do not know the inclination angle α , but we can evaluate a minimum magnetic field strength by setting $\alpha = \pi/2$. In CGS units, for a canonical pulsar with $R = 10^6$ cm, $I_0 = 10^{45}$ g cm², we can approximate the magnetic field strength as $B_0 = 6.4 \times 10^{19} \sqrt{P\dot{P}}$ Gauss. It is conventional, however, to quote the magnetic field at the equator B_s which differs by a factor of 2 (see Lyne and Graham-Smith [110] pg. 71)

such that

$$B_s = 3.2 \times 10^{19} \sqrt{P\dot{P}} \text{ Gauss.} \quad (1.13)$$

In this section, we have introduced some of the simple results that can be obtained by modelling the time evolution of pulsars with a power law. Many advancements can be made on this model, such the existence of a magnetosphere predicted by Goldreich and Julian [70], but in practice this model is consistent with most pulsar observations and provides a useful way to categorise them via their spin-down age and magnetic field. We will frequently refer back to this model as it is a useful platform from which to begin understanding neutron stars.

1.5 Radio pulsar population statistics

Radio pulsars make up the majority of the observed neutron star population and will be the focus of discussion in this thesis. In this section, we will provide some simple empirical population statistics for the normal radio pulsar population: we ignore the millisecond population since they are disjoint from the normal population and have a distinct history, but include the young pulsars. More detailed Monte Carlo based population synthesis studies have been performed by Faucher-Giguere and Kaspi [60] and Popov et al. [138] in which substantive models were made which allow inferences to be performed for the underlying population distribution of all neutron stars. In this section, we provide only summaries of the observations and will discuss their results when relevant. All data in this section is taken from the ATNF pulsar catalogue Manchester et al. [117] and it should be stated that in each case the observed property is an average over all observations made for each pulsar.

For each observable property of the population of neutron stars (such as the frequency), we will present the data as a histogram choosing an appropriate binning size in each instance. In order to make simple inferences about the population, we will also calculate the mean and standard-deviation. We will test normality using the Jones et al. [92] implementation of the d'Agostino [51] test. This results in a *p*-value, which, if less than 0.05, rejects the hypothesis that the data is normal with 95% confidence. We will give this *p*-value in the legend for each observable property and show the normal distribution with the calculated mean and standard-deviation.

In Figure 1.5 we present the data for the three timing properties measured directly from the pulsar timing models. For normal radio pulsars the pulsation frequency, ν , can always be accurately measured provided at least one observation has been made. Several precise observations of a pulsar must be made in order to measure the higher order derivatives of the frequency. As a result, the pulsar catalogue contains missing information and the number of data points for $\dot{\nu}$ and $\ddot{\nu}$ is smaller than the total observed number of pulsars: the exact numbers are given in the caption. By eye, the histograms are clustered and appear to be approximately normal. However, for all three properties, the normal hypothesis is rejected; we note that the level of rejection is dependent on the number of data points. This rejection is not surprising given the complicated physics

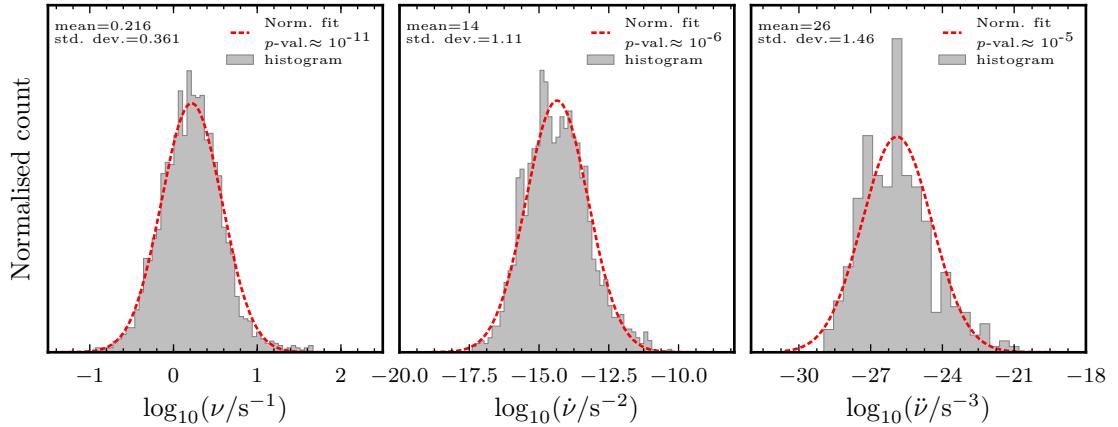


Figure 1.5: The distribution in log-space of the frequency ν and the first two frequency derivatives $\dot{\nu}$ and $\ddot{\nu}$ for normal radio pulsars in the ATNF pulsar catalogue. Appropriate bin sizes were selected for each quantity. The population sizes are 1942, 1686, 339 for ν , $\dot{\nu}$, $\ddot{\nu}$ respectively.

which governs these systems. The detailed population synthesis studies by Faucher-Giguere and Kaspi [60] and Popov et al. [138] are able to relate the observed features to the underlying physics and find similar results for the period and period derivative.

In Figure 1.6 we present some other interesting quantities held in the ATNF catalogue. Firstly, in the left-hand panel we plot the characteristic age as defined in Eqn. (1.10). Then, in the middle panel we give a measure of the pulsar beam-width W_{10} . Specifically, W_{10} is the width of the integrated pulse profile (in seconds) at 10% of the integrated pulse profile maximum. In the right-hand panel we plot $W_{10}\nu$, i.e. the product of the beam-width and frequency for each pulsar. This gives information about the effective duty-cycle: the ratio between the pulse duration and period. Notably, the majority of pulsars have duty-cycles substantially less than a 0.5 indicating that the pulses are short compared to the period. In this instance, the normal hypothesis is rejected for the beam-width and duty-cycle, but accepted for the spin-down age. It would be an interesting exercise to investigate this further.

These results discussed in this section provide an overview of the observed radio pulsar population. It must be remembered that these observed pulsars are a sample from what may be a much larger population. These summaries are intended only to give a brief overview of the observations.

1.6 Neutron stars and gravitational waves

Gravitational waves (GWs) were first predicted by Albert Einstein in 1916 [58] when he found that the linearised weak-field equations of his General Theory of Relativity had transverse wave solutions. Much like the generation of electromagnetic waves requires

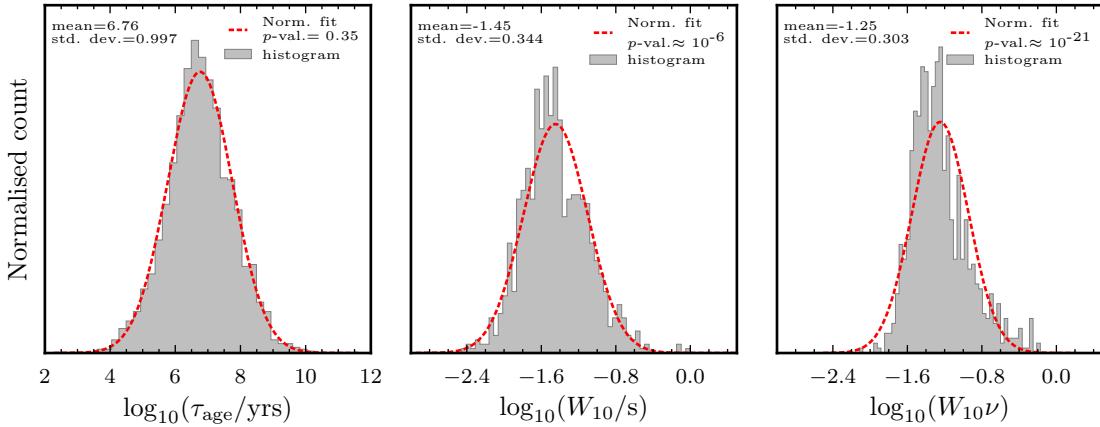


Figure 1.6: The distribution in log-space of the characteristic age τ_{age} , the W_{10} measure of the beam-width, and the effective duty-cycle $W_{10}\nu$ for normal radio pulsars in the ATNF pulsar catalogue. Appropriate bin sizes were selected for each quantity. The population sizes are 1942, 915, and 915 for τ_{age} , W_{10} , and $W_{10}\nu$.

the acceleration of electrical charges, GWs are generated by any source with a time-varying mass quadrupole moment and can be understood as ‘ripples’ or spatial strains in the spacetime itself which travel at the speed of light.

Gravitational waves were first directly detected by the LIGO collaboration [12]. They observed a signal consistent with the inspiral and merger event of two $\sim 30M_{\odot}$ black-holes over approximately 0.2 s. To detect such signals, LIGO uses a *laser interferometer* to measures the relative change in length between two orthogonal arms. In particular, if L is the length of either arm without a signal and a gravitational wave passes through, the detector measures the strain

$$h(t) = \frac{\delta L_x - \delta L_y}{L} \quad (1.14)$$

where δL_x and δL_y are the time-varying stretching and squeezing of the two arms caused by the gravity wave. For the observed binary black-hole merger the peak strain in the detector was $\sim 10^{-21}$.

Prior to this detection, indirect evidence for the existence of gravitational waves was found by observing the orbital periods of compact binary systems. Such systems have a time-varying quadrupole moment and emit gravitational waves, which radiate energy away from the system causing a decay of the orbital period. In 1975, Hulse & Taylor discovered a binary neutron star system where one of the stars, PSR B1913+16, was visible as a pulsar [83]. Due to the powerful techniques of pulsar timing, subsequent analysis by Taylor and Weisberg [159] was able to verify that the orbital decay matched exactly the predictions of General Relativity. Since this observation, more double neutron star system have been discovered, including a system, PSR J0737-3039A/B, discovered by Burgay et al. [37], where both neutron stars are seen as pulsars. This so-called double pulsar system tests the agreement with General Relativity at the 0.05% level [98].

Neutron stars observed as pulsars are often referred to as ‘cosmic clocks’ for the regularity of their pulsations. The most stable pulsars are the radio millisecond pulsars (MSPs), which, due to their stability, many workers in the field utilise in an attempt to search for GWs via a *pulsar timing array* [79]: this searches for correlated signatures in the TOAs from a network of well-timed MSPs. Such a detector is sensitive to a stochastic background of gravitational waves by measuring the so-called Hellings & Downs curve [77], or to the mergers of super-massive black hole binary systems [103].

Isolated neutron stars themselves are potential sources of gravitational waves through one of three mechanisms. If the star has a rotation axis misaligned with its symmetry axis then it will undergo *precession*: a ‘wobble’ of the star which has a time-varying quadrupole moment. This will produce GWs at the rotation frequency and twice the rotation frequency, but the small amplitudes of possible sources and questions over how long lived they might be make this an unlikely candidate for LIGO [91]. If the neutron star is subject to *non-axisymmetric instabilities*, such as the r-mode instability in newborn and rapidly accreting neutron stars [19], then these too can produce GWs (for a review see Andersson [17]). Finally, if the star possesses a *non-axisymmetric distortion*, ϵ , also known as a ‘mountain’, it will produce a continuous gravitational wave at twice its rotation frequency with a strain amplitude proportional to ϵ . The LIGO detectors have already been used to search for signals from known neutron stars and, by not observing any radiation, are able to place upper limits on ϵ (see for example Abbott et al. [11], Abadie et al. [9]).

All three of these detection mechanisms are potential sources of the first detection of gravitational waves from neutron stars and realising this would provide a unique opportunity to learn about neutron stars. But is it feasible? A statistical argument can be made for the ‘loudest expected signal from unknown isolated neutron stars’. This argument is given in Abbott et al. [10], although the origin can be dated back to Blandford (1984) as attributed by Kip Thorne in Hawking and Israel [75]. Essentially, one makes the assumption that the population of 10^5 neutron stars predicted to exist in our galaxy by stellar evolution models are all born with a high spin-down rate and subsequently spin-down principally due to the emission of gravitational waves. With additional assumptions that the stars are born randomly throughout the Galactic disk with a constant birthrate the populations are transformed into a population of neutron star strains. Then it is shown that there is a 50% chance a source exists with a strain amplitude

$$h_0 \sim 4 \times 10^{-24}, \quad (1.15)$$

which is close to ‘detectable’ by LIGO, although the exact details depend on the source frequency and duration. While this is a purely statistical argument, and changing any of the assumptions tends to decrease this signal strain [140], the rewards for detection in terms of astrophysics are sufficient to motivate further research.

1.7 Plan of the thesis

Following this introductory chapter, we will introduce timing variations in pulsars: glitches and timing noise. This will familiarise the reader with the observed phenomena and describe the current state of modelling. We will provide some original work on simple ways in which the models could be tested.

In the next four chapters we evaluate models of timing noise in the face of current observations and attempt to constrain the models. In Chapter 3 we explore how *precession*, a potential ingredient to explain timing noise, can be described, when viewed from the frame rotating with the star. Following this, Chapter 4 looks at how precession will manifest in the observations made by pulsar astronomers. In Chapter 5 we perform a rigorous quantitative model comparison between precession and the leading alternative, *magnetospheric switching*, for describing timing variations seen in PSR B1828-11.

In the final three chapters we approach another important aspect of timing variations for neutron stars: the effect they may have on our ability to detect gravitational waves from neutron stars. In Chapter 6 we introduce the methods and formalisms used by gravitational wave astronomers before analysing the effect of glitches on gravitational wave searches in Chapter 7. For the effect of timing noise on gravitational wave searches, we perform a numerical study on data from the Crab pulsar in Chapter 8 and then model the effect of different timing noise interpretations in Chapter 9. Finally, we will conclude in Chapter 10.

Chapter 2

Timing variations

Timing variations are, broadly speaking, any time when the usual Taylor expansion in the phase Eqn. (1.2) (typically up to the second-order spin-down rate $\ddot{\nu}$) does not accurately describe the phase evolution of the pulsations. Two distinct types of variations exist: the sporadic event-like *glitches* seen in some pulsars and the ubiquitous *timing noise* present, at some level, in all pulsars. In this chapter, we will discuss the observations of these variations and the models proposed to understand them.

2.1 Glitches

In addition to the regular spin-down of isolated radio pulsars due to magnetic braking, some pulsars undergo anomalies in their timing solutions known as *glitches*. These are sudden rapid increases* in the pulsation frequency which were first observed in the Crab [34, 147] and Vela pulsars [144, 145]. Pulsar timing methods model this as a permanent increase in the phase, frequency, and first frequency derivative in addition to a frequency increment that subsequently decays exponentially to zero. To model this, for each glitch pulsar astronomers add on an additional term to Eqn. (1.2) which is Eqn. (121) of Edwards et al. [57][†]:

$$\phi_g = H(t - t_g) \left(\Delta\phi + \Delta\nu(t - t_g) + \frac{\Delta\dot{\nu}}{2}(t - t_g)^2 + \left[1 - \exp\left(-\frac{t - t_g}{\tau}\right) \right] \Delta\nu_t \tau \right), \quad (2.1)$$

where $H(t)$ is the Heaviside step function. The first three terms are the permanent increase in phase, frequency, and spin-down, while the last term gives the transient increase in the frequency $\Delta\nu_t$ which decays exponentially with a timescale τ . To illustrate this, in Figure 2.1 we show the spin-frequency model of a glitch including a permanent increase in frequency $\Delta\nu$ and a component $\Delta\nu_t$ which is ‘recovered’.

*A single ‘anti-glitch’ in an isolated magnetar has also been reported by Archibald et al. [21] in which the pulsation frequency spontaneously decreased. The implications of this remain unclear, but its existence was further confirmed by Hu et al. [82].

[†]Note that there is a typographical error in Eqn. (121) of Edwards et al. [57]: the final $t - t_g$ should in fact be τ .

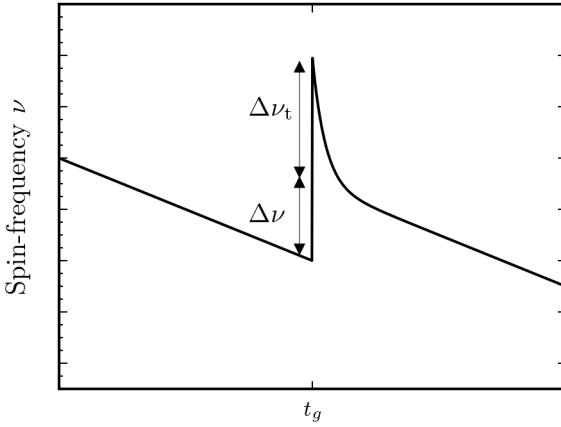


Figure 2.1: Illustration of the glitch model fitted by pulsar astronomers.

In effect, pulsar astronomers fit separate Taylor expansions either side of the glitch. This is a good model when the rise-time of the glitch, during which the frequency increases, is short compared to the duration between observations. This evolution of the frequency during a glitch has yet to be observed, but high time resolution monitoring of the Vela pulsar placed an upper limit of 40 s for the rise-time between the original and the new period [55]. Since we cannot resolve the glitch itself, Eqn. (2.1) is appropriate and used for all known glitches.

A comprehensive review of glitches was carried out by Espinoza et al. [59]; to illustrate a typical glitch, in Figure 2.2 we reproduce data from this review on a glitch in the Crab pulsar.

Over 165 of the \sim 2000 observed pulsars have been seen to glitch, often multiple times. Typical values of the instantaneous frequency change range from 10^{-9} Hz to 10^{-4} Hz. For some pulsars this is accompanied by a change, with either sign, in the spin-down rate $\Delta\dot{\nu}$ with absolute magnitudes between 10^{-19} Hz/s to 10^{-12} Hz/s. The *glitch recovery parameter* is defined as

$$Q = \frac{\Delta\nu_t}{\Delta\nu + \Delta\nu_t}. \quad (2.2)$$

Most glitches are not resolved with sufficient detail to determine this recovery parameter accurately. A review of those pulsar glitches with measured values of Q was conducted by Lyne et al. [114]; they found that in glitches from 18 pulsars, Q correlates with $|\dot{\nu}|$ reaching values as large as ~ 0.9 for the youngest pulsar with the highest absolute spin-down rate, the Crab pulsar.

Many pulsars have been observed to glitch several times, Melatos et al. [123] considered the waiting times between glitches and concluded that in most glitching pulsars the glitches happen randomly with waiting times consistent with a Poisson process, except in PSR J0537-6910 and PSR B0833-45 (the Vela pulsar) which displayed quasi-periodicity in the waiting times.

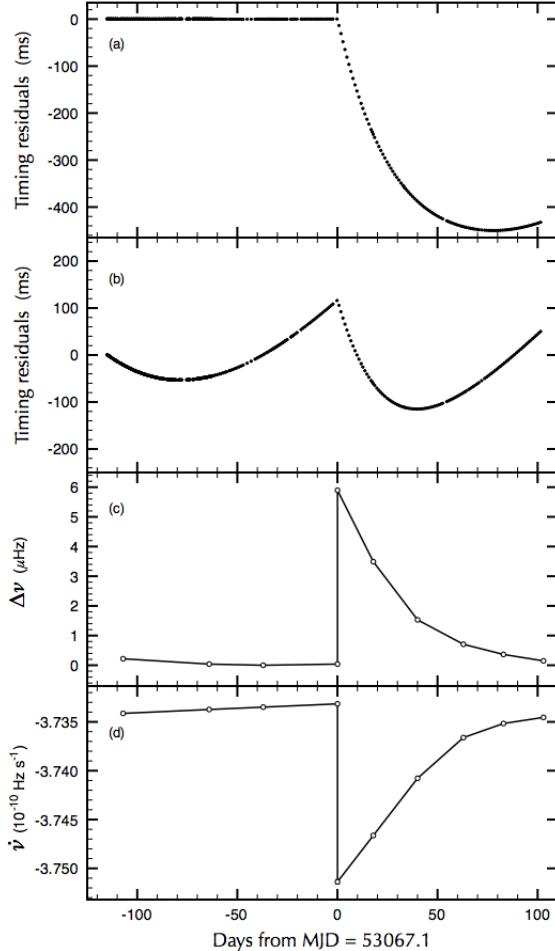


Figure 2.2: A glitch in PSR B0531+21, the Crab pulsar. It occurred around MJD[‡] 53067 and had a permanent fractional frequency jump of $\Delta\nu/\nu = 5.33 \pm 0.05 \times 10^{-9}$. (a) The timing residuals relative to a slowdown model with two frequency derivatives when fitting data only up to the glitch date. (b) Timing residuals after fitting all data in the plot; note that the glitch feature is still visible. Both these panels have the same y -scale, covering 500 ms. (c) Frequency residuals, obtained by subtracting the main slope given by an average $\dot{\nu}$. (d) The behaviour of $\dot{\nu}$ through the glitch. This figure and caption are adapted from Figure (1) of Espinoza et al. [59].

In Chapter 7, we perform our own investigation into the population statistics of glitches with an aim to understand their implication for gravitational wave searches. We find, in agreement with Espinoza et al. [59] and references therein, that the distribution of glitch magnitudes has multiple modes which suggests that glitches may come from more than one mechanism. We go on to apply a statistical model and determine empirically the properties of the underlying source populations.

Glitches provide a unique opportunity to investigate the physics of neutron stars. The two models able to explain some of the observations are known as the *superfluid unpinning* model and the *starquake* model.

[‡] The modified Julian date (MJD) calendar counts days from an epoch of midnight on the 17th of November 1858.

In the superfluid unpinning model proposed by Anderson and Itoh [16], the star contains a superfluid component in which the angular momentum is stored in an array of vortices which are ‘pinned’ to the crust. The magnetic dipole, rigidly fixed to the crust, exerts a torque on the crust gradually spinning it down. The superfluid component cannot decrease its angular momentum without reducing the number of vortices per unit area, so does not spin down at the same rate. A lag in frequency between the superfluid component and rest of the crust develops until the forces are sufficiently large to cause an avalanche of unpinning events rapidly transferring the stored angular momentum in the superfluid component to the crust. An observer measures the frequency and spin-down rate from the rate of pulsations. Since these pulsations originate from the EM dipole which is frozen into the crust of the star, when this unpinning occurs, we see a rapid increase in the frequency.

The second model, starquakes, follows from the observation that a rapidly spinning fluid body has an oblate ‘rest shape’ with a bulge about its equator due to the centrifugal force. The crust of a star spinning at some frequency will solidify as the star cools, with a corresponding oblateness which we call the reference oblateness. Subsequently, as the star spins down, it will have a different rest oblateness due to its decreased frequency, but the crust will retain a memory of the earlier reference oblateness at which it solidified. This will cause strains in the crust which eventually cause a starquake relieving the strain, resetting the reference rest shape, and producing glitch like features. This model was first proposed by Ruderman [148] and later built upon by Baym and Pines [28].

Both of these models have support in the literature and have been developed significantly to explain the variety of observed glitches. However, there are observations which cause difficulties for both models: glitches seen in the Vela pulsar are too large and too often to be consistent with a starquake model [39], while the unpinning model requires a superfluid component which is at odds with observation of precession (we discuss this further in Chapter 5). In this thesis, we will not use glitches as a tool for inferring neutron star physics, but any predictions we do make must be compatible with what has already been learnt from glitches.

2.2 Timing noise: observations

Timing noise refers to small-scale structure in the timing residual which cannot be attributed to any other source and hence cannot be modelled and included in the timing model. The presence of timing noise indicates that we do not have a complete picture of the neutron star: there is unmodelled physics.

Characterising timing noise is a difficult task: the exact form it takes will depend on the order of Taylor expansion used to fit the timing parameters. Typically, pulsar astronomers truncate at $\ddot{\nu}$, but fitting to higher orders is possible and will tend to decrease the ‘level’ of the resulting structure understood as timing noise. Of course using a sufficiently large number of terms in the Taylor expansion, eventually one will fit out all the structure. However, this does not provide any additional insight into the

cause of timing noise. In this thesis, we will define timing noise as the remainder having fitted and subtracted a second order Taylor expansion. In the literature, a second order fit is most commonly used (see for example Hobbs et al. [81]), but examples exist of fitting and removing higher order Taylor expansions (e.g. PSRs B0919-06, B1540-06, and B1828-11 in Lyne et al. [111]).

To illustrate timing noise and how it can depend on the order of Taylor expansion used, in Figure 2.3 we show the phase residual remaining after fitting and removing a 3rd, 4th, and 5th order Taylor expansion to the Crab pulsar. In all three instances we see a quasi-periodic structure reaching residuals up to half a cycle; higher order have lower residuals, but the form of the structure remains consistent between orders.

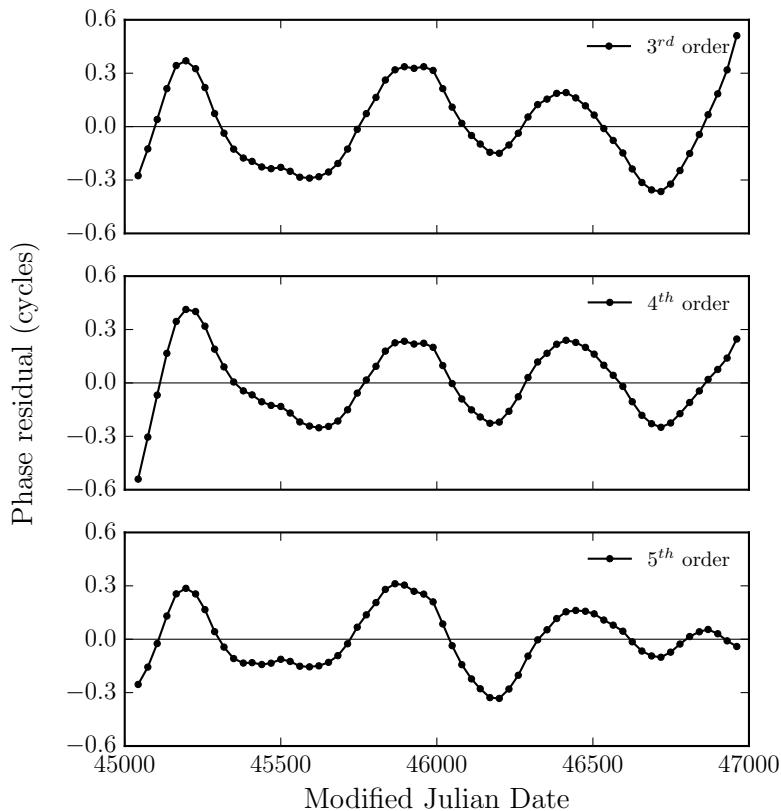


Figure 2.3: A phase residual demonstrating the structure which is named timing noise. This is generated from data on the Crab pulsar (see Section 8.2 for details).

Several methods exist in the literature to quantify the strength of timing noise such as the Δ_8 value introduced by Arzoumanian et al. [23], the generalisation of the Allan variance [119], the covariance function of the residuals [42], and fitting for timing noise as part of the pulsar timing model [104]. The most comprehensive and recent analysis was performed by Hobbs et al. [81] who considered 366 pulsars over timescales $\gtrsim 10$ years. We summarise their conclusions here:

1. Timing noise is widespread in pulsars
2. Timing noise is inversely correlated to the characteristic age as defined in Eqn. (1.9)
3. The structures seen in the timing residual vary with data span: as more data is collected, more quasi-periodic features are observed.
4. The dominant contribution to timing noise for young pulsars with $\tau_{\text{age}} < 10^5$ years can be explained as being caused by the recovery from previous glitches.
5. A handful of pulsars exhibit significant periodicity while quasi-periodicities are observed in many pulsars

These general features give a broad picture, but there is great variation in the form of timing noise between pulsars; this is illustrated by the variety of timing residuals reported in Hobbs et al. [81]. To understand the variety of observation, in the next section we will discuss some of the models for timing noise that exist in the literature and which observations they are able to explain.

2.3 Timing noise: interpretations

The underlying mechanism which causes timing noise is not understood. Since the first discussions in Boynton et al. [33] multiple models have been proposed which are able to describe some of the features. However, the variety of ways timing noise manifests and the uncertainty in the mechanisms at work have made it difficult for any conclusive statements to be made about the models. A complete understanding of timing noise must not only explain the observed variations, but also remain consistent with our understanding of neutron stars derived from other observations such as glitches. A complicating factor in understanding timing noise is that the observed features have timescales similar to the duration that we have been able to observe pulsars; it is therefore possible, as noted by Hobbs et al. [81], that observations which look like a random walk over a short timescale, may in fact be periodic or something else entirely over longer timescales. Observations of new features prompt new models for timing noise; as a result the timing noise interpretations have evolved with the observations. In this section, we will present an overview of these interpretations and the evidence which supports them.

2.3.1 Random walk models

Timing noise was first quantified and interpreted by Boynton et al. [33] in the context of the Crab pulsar as a Poisson like random walk in one of the phase, frequency, or spin-down. The pulsar spins down according to the power law spin-down of Eqn. (1.7) except that at random times the pulse phase, frequency, or spin-down will undergo a sudden step; these are referred to as phase noise (PN), frequency noise (FN) and spin-down noise (SN). The waiting times between steps are Poisson distributed with a rate λ . The number of events in such a model is considered to be large such that in a

typical observation time T_{obs} we observe the superposition of many events. Following the method of Cordes [43], the *strength* of the three types of timing noise can be defined as

$$S_{\text{PN}} = \lambda \langle \delta\varphi^2 \rangle, \quad (2.3)$$

$$S_{\text{FN}} = \lambda \langle \delta\nu^2 \rangle, \quad (2.4)$$

$$S_{\text{SN}} = \lambda \langle \delta\nu^2 \rangle, \quad (2.5)$$

where $\langle \delta\varphi^2 \rangle$, $\langle \delta\nu^2 \rangle$, and $\langle \delta\nu^2 \rangle$ are the mean-square values of the steps in phase, frequency, and spin-down rate. These strength parameters form the basis of much of the analysis of random walk models. They are determined by separating the observed TOA into blocks of data of duration T and fitting a timing model to each block. In each block, one can then calculate the phase residual $\Delta\varphi(t)$. Time averaging the squared phase residual for each block of duration T and then averaging over all the blocks of data, we define

$$\sigma_\varphi^2(T) = \left\langle \frac{1}{T} \int_0^T \Delta\varphi(t)^2 dt \right\rangle, \quad (2.6)$$

which is an estimate of the variance of the phase residual over a time span T . Cordes [43] then showed that this is related the strengths of each noise component by

$$\sigma_\varphi^2(T) = \begin{cases} \frac{S_{\text{PN}} T}{2}, \\ \frac{S_{\text{FN}} T^3}{12}, \\ \frac{S_{\text{SN}} T^5}{120}. \end{cases} \quad (2.7)$$

Combining this with an estimate of σ_φ^2 from the data using Eqn. (2.6), these relations give the strength of each noise component for blocks of duration T .

Given a sufficient amount of data, one can then calculate the noise strengths for different values of T . If the strengths vary with the T , the random walk component can be rejected since this is not predicted by the model. However, if the strength for a given noise component is robust to changes in T , this provides evidence that the random walk is in fact occurring. However, this consistency is a necessary, but not sufficient condition; it can't be ruled out that another mechanism mimics this behaviour. Furthermore, the situation can be complicated by mixing of noise processes as shown in Groth [73] and Cordes [43].

Using a similar process, Boynton et al. [33] categorised timing noise in the Crab pulsar as frequency-like. They found that over a 5 year period the noise process was stationary and consistent with the frequency noise hypothesis. No deterministic process could account for the timing residuals strengthening their conviction that some random process was taking place. The Crab was subsequently analysed by Groth [73] and Cordes [43]; the analysis was extended to other pulsars by Helfand et al. [76] and Cordes and Helfand [48].

The model of timing noise as a Poisson random walk is a purely empirical statistical one. It is however backed up for a rich variety of possible substantive physical models. A key feature of any physical random walk model is that it must be able to produce both

increases and decreases in the relevant parameter. For this reason, it is unlikely that the mechanism responsible for a random walk timing noise is the same as that proposed to explain glitches. Since it is assumed that the number of events is large so that we can take a statistical average, this does not test if the process is indeed discrete, or if it is continuous.

The first substantive physical model to explain the random walk was proposed by Boynton et al. [33]. The noise process consisted of the accretion of small lumps of matter onto the star from the interstellar medium. Lumps of matter fall randomly onto the surface of the star causing either a spin-up or spin-down through the transfer of angular momentum. After this, many models were proposed such as small starquakes and the random pinning and unpinning of vortex lines; these were reviewed by Cordes and Greenstein [47] and evaluated against observational constraints. Of these, only three mechanisms were found to be consistent with observations: crust breaking by vortex pinning, a response to heat pulses, and luminosity related torque fluctuations. Since this review, new random walk mechanisms have been proposed such as variations in the magnetospheric gap size [40], the interference by debris entering the magnetosphere [49], and the accumulation of multiple micro-glitches [84]. It would be a useful exercise to compare both the old and new mechanisms against the current observational catalogue since improvements in the quality and duration of data and a larger number of sources may better constrain some of these physical models.

The first measurement of apparent individual timing noise events was made by Cordes and Downs [46] who identified ~ 20 step-like events in both frequency and spin-down which could not be explained by a glitch. By considering 24 pulsars over a period of ~ 13 years, the authors concluded that the timing noise seen in the data could not be explained solely by an idealised random walk processes in either the phase, or its derivatives. They suggested instead that most of the activity is due to a mixture of events in the phase, frequency and/or frequency derivative.

A recent observational review of timing noise was performed by Hobbs et al. [81] for 366 pulsars. The authors do not use the observations to constrain random walk models. Instead, they state that timing noise cannot be explained by a simple random walk model and when observed on sufficiently long timescales, the residuals which may before have looked like a random walk, contained quasi-periodic features. It is impossible to argue that this is not the case for pulsars which only display random-walk features over current observation periods, since the quasi-periodicity may have periods longer than the observation. This effect is apparent for the Crab pulsar which was found by Boynton et al. [33] to be consistent with frequency-like noise, but in Figure 2.3 we see quasiperiodic features. On the basis of Figure 2.3, we agree that it is plausible that there are features of the Crab timing that require more than a random walk. However, if one simply wants to characterise timing noise over short periods of data, a random walk provides a good empirical description.

2.3.2 Free precession

A mechanism which could quite naturally produce strictly periodic variations in the observable features of a pulsar is *free precession*. This occurs in any non-spherical rigid body for which the angular momentum is not aligned with a principal axis of the moment of inertia. Such a circumstance could arise given the chaotic birth of neutron stars. However, we must be clear that the timing noise induced by precession alone would be strictly deterministic; this is something which we do not typically observe except in a handful of pulsars, one of which is considered in Chapter 5. We will now consider the mechanics of free precession.

In the simplest case, we take a biaxial body, rotating about an axis Ω , with a moment of inertia given by

$$I = \begin{bmatrix} I_0 & 0 & 0 \\ 0 & I_0 & 0 \\ 0 & 0 & I_0 + \Delta I \end{bmatrix} = \begin{bmatrix} I_0 & 0 & 0 \\ 0 & I_0 & 0 \\ 0 & 0 & I_0(1 + \epsilon_I) \end{bmatrix}, \quad (2.8)$$

where I_0 is the total moment of inertia of the star involved in precession and $\epsilon_I = \Delta I/I_0$ is the measure of oblateness or prolateness which participates in precession. This asymmetry ΔI , refers only to the portion of the deformation that follows the precession of the star, i.e. that part source by strain, not the part sourced by the centrifugal deformation Jones and Andersson [90]. If the body is free from torques, then in the rotating frame of the body, Euler's equations of motion [100] are given by

$$I\dot{\boldsymbol{\Omega}} + \boldsymbol{\Omega} \times (I\boldsymbol{\Omega}) = 0. \quad (2.9)$$

This is a system of three coupled ordinary differential equations (ODEs). Writing the components of the spin-vector as $\boldsymbol{\Omega} = [\Omega_x, \Omega_y, \Omega_z]$, we have

$$\dot{\Omega}_x = -\epsilon_I \Omega_y \Omega_z, \quad \dot{\Omega}_y = \epsilon_I \Omega_x \Omega_z, \quad \dot{\Omega}_z = 0 \quad (2.10)$$

We can find a solution by first setting $\Omega_z = \text{const}$. We are then left with a set of two coupled ODEs, solving these with appropriate initial conditions the solutions take the form

$$\Omega_x = \Omega_0 \sin(a_0) \sin(\Omega_0 \cos(a_0) \epsilon_I t), \quad (2.11)$$

$$\Omega_y = \Omega_0 \sin(a_0) \cos(\Omega_0 \cos(a_0) \epsilon_I t), \quad (2.12)$$

$$\Omega_z = \Omega_0 \cos(a_0), \quad (2.13)$$

where a_0 is the angle between the spin-vector and the body frame z axis, and Ω_0 is the magnitude of the spin-vector.

We observe that the spin axis of the body will trace out a cone about the z principal axis of the moment of inertia with a period of

$$\tau_P = \frac{2\pi}{\Omega_z \epsilon_I}. \quad (2.14)$$

The half-angle of the cone is set by the initial conditions and will not evolve. This is the motion of free precession and is illustrated in Figure 2.4.

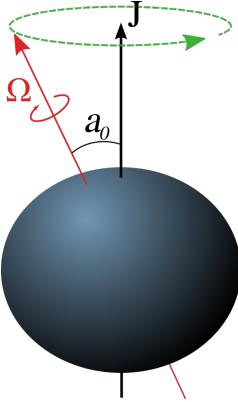


Figure 2.4: Illustration of free precession for a simple biaxial body. The spin axis Ω traces out a cone about the angular momentum vector \mathbf{J} .

It was shown by Ruderman [148] that neutron stars should have a solid crystalline crust which is capable of supporting shear stresses such that there may be differences in the components of associated moment of inertia tensor. Precession, as a candidate to explain timing noise fluctuations, was first discussed by Ruderman [149]. He found that the free precession period was, for an ellipticity of $\epsilon_I \sim 10^{-9}$, able to explain periodic fluctuations in the Crab pulsar. Precession is one of the few periodic mechanism that could act on neutron stars with the periods observed in typical timing residuals.

However, the superfluid unpinning interpretation for glitches poses a problem for sustained free precession as an interpretation of timing noise. If correct, then the interior of a neutron star must contain a superfluid component pinned to the crust. It was shown by Shaham [152] that for perfect pinning the deformation relevant to precession is modified to give

$$\epsilon_I = \frac{\Delta I}{I_0} + \frac{I_{\text{psf}}}{I_0}, \quad (2.15)$$

where I_{psf} is the moment of inertia of the superfluid pinned to the crust. To explain the magnitude of observed large glitches this second term is typically found to be ~ 0.01 (see for example Table (1) of Andersson et al. [18]). Inserting this into Eqn. (2.14), we see that the free precession period must be approximately 100 times the spin-period. Therefore, the existence of a perfectly pinned superfluid precludes long-period precession where the precession period is millions of times longer than the spin-period.

Despite the inconsistency with the superfluid pinning model for glitches, evidence was presented by Stairs et al. [156] of free precession in PSR B1828-11. They found the phase residuals and variations in the pulse profile could be accounted for by precession. This pulsar has been studied by others since and in Chapter 5 we will discuss this pulsar and the precession interpretation in more detail.

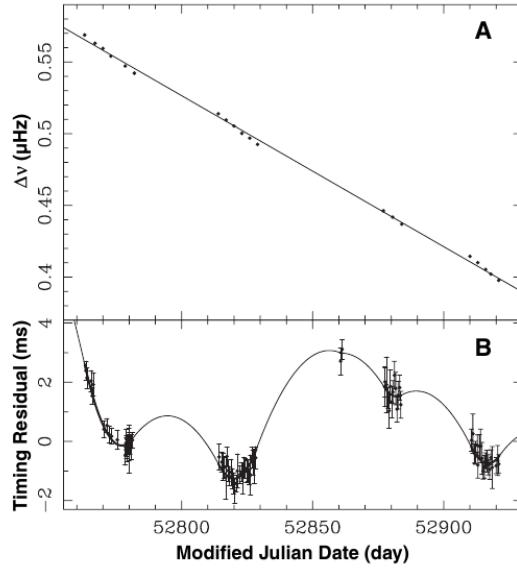


Figure 2.5: Figure taken from Kramer et al. [97] showing the switched spin-down of pulsar PSR B1931+24

One resolution to this comes from relaxing the assumption that the superfluid pinning is perfect. Sedrakian et al. [150] found that imperfect pinning of the superfluid allows long-lived precession with damping. It is therefore possible, as noted by Cordes [44] that free precession may be excited by torque fluctuations that counter the damping process; in turn, the precession can drive torque fluctuations. However, there remains some uncertainty as to the details of this model, so whether precession can coexist with a superfluid core remains an open question. Resolving this, and determining exactly how the models are incompatible is an important task which we hope to make some headway on in this thesis by studying the case for precession in Chapter 5.

2.3.3 Magnetospheric switching

Recently a new model has been proposed by Lyne et al. [111] to explain the observation of quasi-periodicity in timing noise structure. This work was motivated by the observations of Kramer et al. [97] that the pulses from PSR B1931+24 were intermittent: the pulsar acts as a normal pulsar for ~ 10 days and then switches off, being undetectable for ~ 25 days, before switching on again. This behaviour can be understood as the pulsar switching between two states. Analysing the spin-down rate between the on and off states, they determined the spin-down rate $\dot{\nu}$ was $\sim 50\%$ faster in the on state. The figure illustrating this is reproduced in Figure 2.5. The upper panel, Figure 2.5A, shows the evolution of the rotational frequency over a 160 day period encompassing several switching events. A line shows the long-term averaged spin-down of the pulsar while the dots show individual measurements made during the on state. During these on states, the gradient of the reduction in frequency is increased: the spin-down rate is faster. It is thought that measurements of the frequency in the off state would produce a line with

decreased spin-down connecting the dots. Plotted in Figure 2.5B is the timing residual; this shows significant quasi-periodic modulations correlated with the switching.

PSR B1931+24 evidently switches suddenly and periodically between two distinct states. Kramer et al. [97] realised that the neutron star magnetosphere is the only mechanism which could produce such sharp changes with correlated changes in the spin-down rate and beam. It makes sense that magnetospheric state which does not produce EM emission may have a lower spin-down rate while the state which does produce radiation has a higher spin-down rate.

Motivated by this observation, Lyne et al. [111] tested a set of other pulsars and presented a study of 17 pulsars for which they found evidence for ‘two-state magnetospheric switching’. Unlike in PSR B1931+24, these pulsars are not intermittent, but are seen to continuously pulse. The authors measured changes in the spin-down rate $\dot{\nu}$ of each pulsar by using what we will call in this thesis the *observer-method*. This consists of fitting a Taylor expansion to short segments of data of duration T . In each segment, the fitted coefficient $\dot{\nu}$ is taken as a measurement of the spin-down rate at the mid-point of the segment. Repeating this process in a sliding window at intervals $T/4$ throughout the whole data set gives the evolution of the spin-down rate over the entire observation. Over a ~ 20 year period they found a variety of smooth periodic fluctuations with typical periods of years. In Figure 2.6 we reproduce their original plot showing the periodic variations in spin-down rates.

The method used to calculate the spin-down in Figure 2.6 required time averaging; the baseline used T is given below the pulsar name on the right-hand side of the plot, typically $T \sim 100$ days. Lyne et al. [111] argue that, if the pulsar is in fact switching between magnetospheric states, then its spin-down rate will be periodically switching between at least two values. In the event that the switching occurs over timescale comparable to the time averaging baseline, the resulting time averaged spin-down rate will be smoothed out. Therefore, the smooth periodic variations in Figure 2.6 could be produced by the spin-down rate switching between two (or more) well defined values.

If the switching was magnetospheric in origin, then Lyne et al. [111] realised that changes in the spin-down rate should correlate with changes in the beam-shape. They found that for 6 pulsars the pulse width did indeed correlate with changes in the spin-down rate, although for some pulsars this was an anti-correlation rather than a correlation. However, like the spin-down rate these pulse shape variations are also smooth and subject to the same time averaging process.

To test if the mechanism causing these variations is smooth or instantaneous (i.e. switching) the authors needed to look at an observed property not subject to the time-averaging process. This is possible by looking at individual measurements of the pulse width which are taken during each short ~ 30 minute observation. In this way, they are time averaged over a duration much shorter than the modulation period and could, in principle, resolve individual switching events. They were able to do this for two pulsars, PSRs B1822-09 and B1828-11, and argue that the individual beam-width measurements, by eye, can be interpreted as taking one of two values with few observations seen in between. They took this as evidence that the pulsars were indeed undergoing sudden, instantaneous,

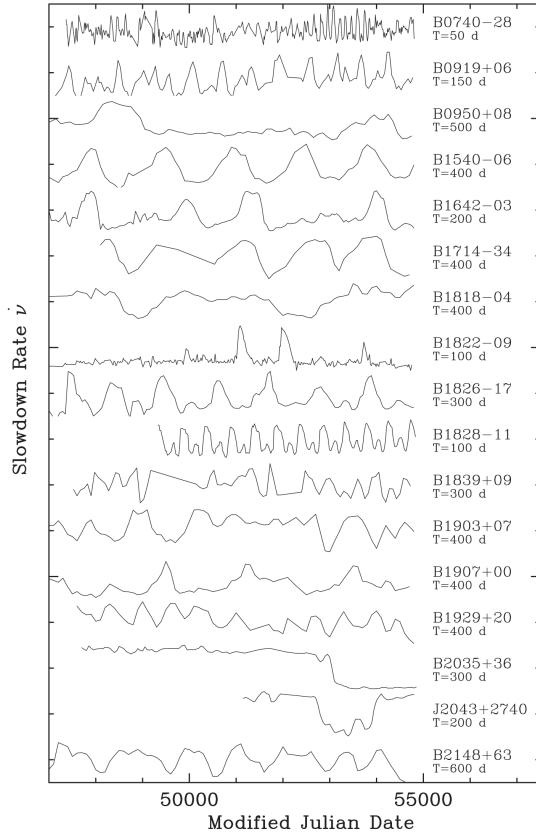


Figure 2.6: Figure adapted from Lyne et al. [111] showing the spin-down rate of 17 pulsars over a ~ 20 year period.

and periodic switching between two states. We note that the stronger of these two candidates, PSR B1828-11, was previously discussed in the context of free precession in Section 2.3.2. Lyne et al. [111] argue that the fast state changes seem to rule out free precession as the origin of oscillatory behaviour observed in timing residuals, in particular for PSR B1828-11.

Since this initial study, evidence for switching in two others pulsars have been reported by Perera et al. [134] and Perera et al. [135] with additions added to the model in terms of the detail of the switching. However, so far the magnetospheric switching hypothesis is missing an underlying physical model which both causes the switching between states and provides the clock regulating the period.

Jones [89] argues that the dismissal of precession is premature since the modulation period of the switching has yet to be explained. Instead, the author raises the idea that precession and magnetospheric switching are not mutually exclusive, but precession may in fact cause the switching. Pulsars are most probably born in a randomly distributed magnetospheric state. At least some may therefore exist under a delicate balance between two states. Precession may be capable of periodically varying the statistical probability of existing in one state or the other. Sharp changes would be caused by an ‘avalanche effect’ as the particle energies reach a threshold. This provides the

timescale for switching along with the ability for the switching to be quasi-periodic since the precession only biases the probability.

Another idea proposed by Cordes [45] interpreted two state switching as evidence for a system in a state of ‘stochastic resonance’. This occurs in systems in which, under certain conditions, a weak periodic forcing function is amplified by stochastic noise. To explain this phenomenon in Appendix 2.A we present a toy model of stochastic resonance for a particle in a well. The switching could therefore be the result of any periodic modulation, such as precession, coupled to random fluctuations. This would quite naturally explain the stability of states, the timescales over which they occur, and the fact that it is observed in only some pulsars.

As already mentioned, determining the cause of periodic modulations has important implications for neutron star physics. Later on in Chapter 5 we will investigate this issue, particularly in the context of PSR B1828-11, using a quantitative model comparison. In the remainder of this section we will build a model for magnetospheric switching incorporating all the ideas of Lyne et al. [111] and Perera et al. [134]; this model forms the basis of the model compared against precession in Chapter 5. Finally, in Section 2.3.3.3 we show how this simple empirical model of switching makes a testable prediction which could be checked by pulsar astronomers using current data.

2.3.3.1 Simple empirical model

In the supplementary material to Lyne et al. [111], the authors presented results from a simple empirical switching model to demonstrate the resulting phase residuals. In this section, we will similarly develop a simple model and show the effect of switching on the time-averaged spin-down rate and phase residuals.

We model a pulsar as spinning down in the usual way, except that its magnetosphere periodically and suddenly switches between two different states which we label A and B . The star spends a time t_A and t_B in each state such that $t_A + t_B$ gives the total period of switching. We then associate spin-down rates, $\dot{\nu}_A$ and $\dot{\nu}_B$ to each of the states such that the underlying spin-down rate is a square wave oscillating between these values. Since we do not concern ourselves in this description with the exact periods, only the gross features, let us define the ratio of time spent state in state A and B as $R = t_B/t_A$.

This is a purely deterministic model and, having generated the spin-down values, we can integrate twice to get the phase. From the phase, we can use the observer-method to calculate the time-averaged spin-down rate $\langle \dot{\nu} \rangle$, or we can fit a polynomial to the entire phase evolution and remove it to get a phase residual. This integration to the phase and then applying the observer-method models the data collection mechanism so that our resulting spin-down rate predictions are smooth.

In Figure 2.7 we show the underlying spin-down rate $\dot{\nu}$, the time-averaged $\langle \dot{\nu} \rangle$, and phase residual (in cycles) for some typical values. For the time-averaged spin-down rate we carefully chose an averaging time similar to the switching period. For the final phase

residual we fit and remove a Taylor expansion up to $\dot{\nu}$ and divide the residual difference by the period in order to get the phase residual in cycles.

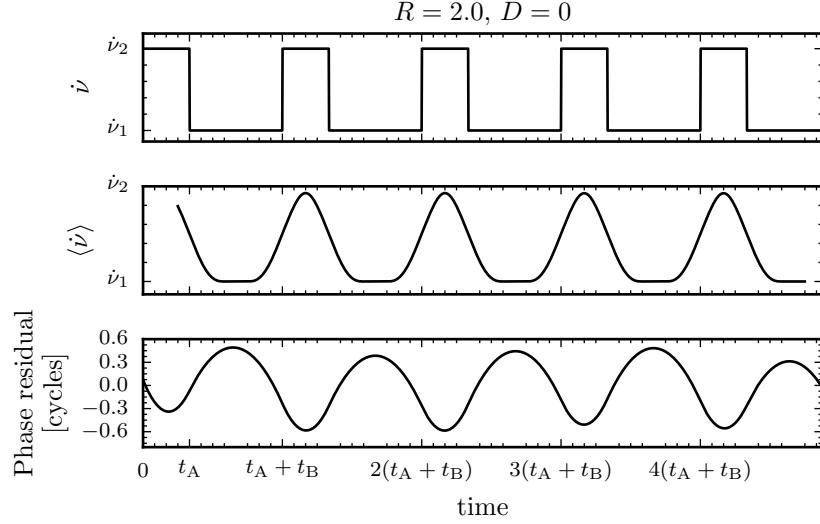


Figure 2.7: A deterministic realisation of the Lyne switched spin-down model. The resulting structure in the time-averaged spin-down rate and the phase residuals are strictly periodic. This shows the underlying spin-down rate in the top panel, the time-averaged spin-down rate (calculated using the observer-method) in the second panel, and the phase residual in the bottom panel.

The phase residuals in Figure 2.7 are strictly periodic which is inconsistent with the observed variations in the majority of pulsars [81]. To develop this, Lyne et al. [111] added a probabilistic 'dither' D in the waiting time between switches. Now we have periods t_A^i and t_B^i which are Gaussian distributed with a mean of t_A and t_B and a standard deviation Dt_A and Dt_B . In Figure 2.8 we repeat the results of Figure 2.7 with $D = 0.3$. This dither can be used to understand why the residuals and averaged spin-down rate are quasi-periodic rather than strictly periodic. However, this model is missing a substantive physical explanation, namely what causes the switching and why it is quasiperiodic.

2.3.3.2 Simple empirical model: four time periods

The time averaged spin-down rates in Figure 2.7 and Figure 2.8 result from having two spin-down rates and two durations which the system spends in those states. These results can be contrasted with the spin-down rates of PSR B1828-11 seen in Figure 2.6. With only the ingredients described by Lyne et al. [111], it is not possible to consistently replicate the double peaked spin-down rate variations seen for this pulsar. A similar problem was faced by Perera et al. [134] for PSR B0919+06, in response the authors devised the following solution. Rather than introduce a third spin-down state, they simply required that the system has four periods instead of two as in the original Lyne et al. [111] description; that is, we have four times t_A, t_B, t_C , and t_D which sum up to

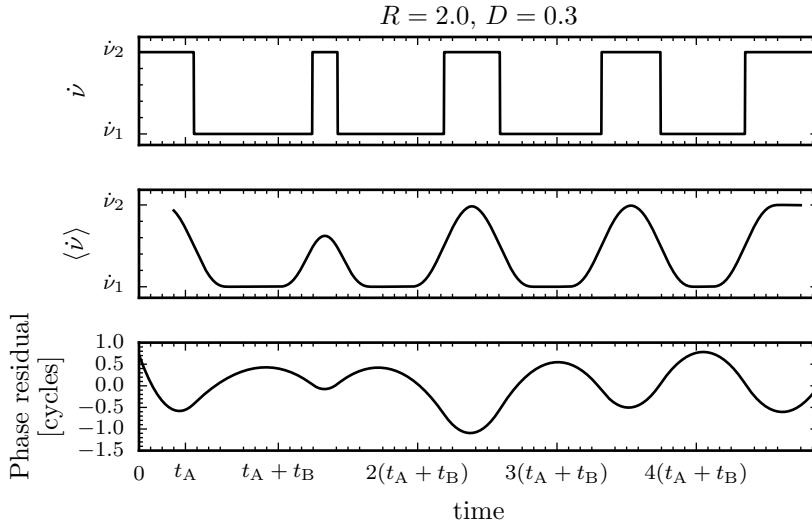


Figure 2.8: A realisation of the Lyne model with a random element producing the observed quasi-period structure. For a description of the three panels see Figure 2.7.

the total period. The system then switches between the two states four times during a single cycle. To illustrate this, in Figure 2.9 we show a single cycle in which the four periods are 100, 100, 50, and 100 days.

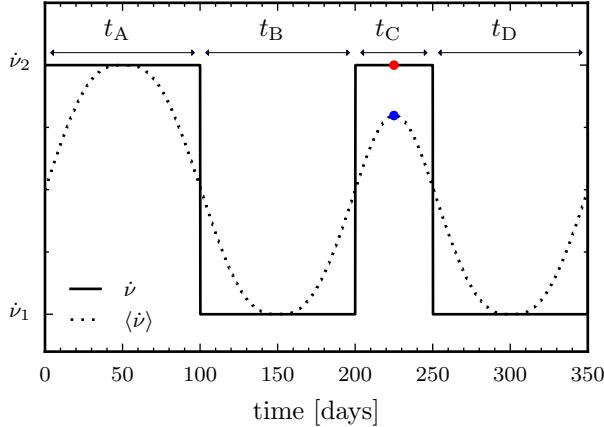


Figure 2.9: The spin-down rate $\dot{\nu}$ and its time-average $\langle \dot{\nu} \rangle$ for a single cycle of the Perera et al. [134] switching model in which the system switches between the two states four times during a single cycle. For the time-averaged spin-down rate, we have used a baseline of 100 days, which is longer than the shortest period of $t_C = 50$ days; for this period we highlight the true spin-down rate by a red dot and the maximum time-average spin-down rate by a blue dot.

The Perera et al. [134] switching model naturally produces the doubly peaked time-averaged spin-down rate seen in PSR B1828-11 when the time-averaging baseline is longer than one of the switching periods. To explain more fully, when the time averaging

baseline is longer than a single period, then the time-average over that section of the data will always include some amount of the other spin-down rate. This can be seen in Figure 2.9 for the third period which has duration 50 days, while the time averaging baseline is 100 days. In Figure 2.10, we give the spin-down rate and phase residual for this model to illustrate the variations in the residuals.

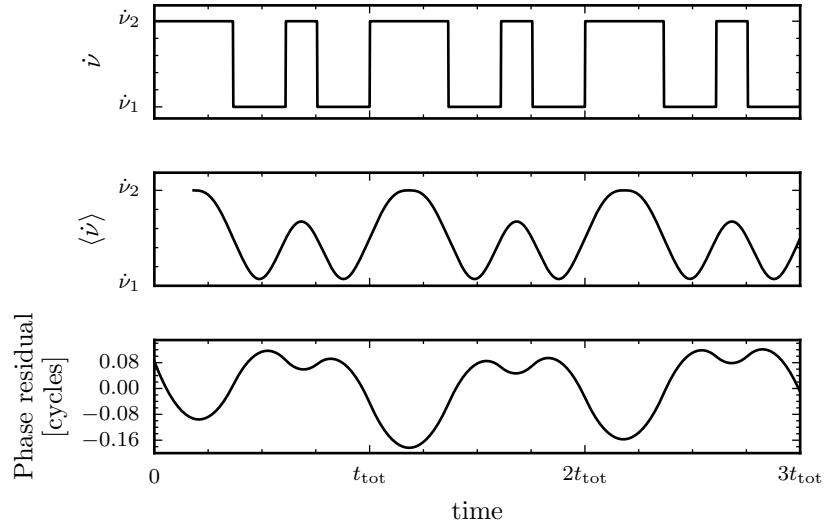


Figure 2.10: Illustration of the Perera et al. [134] extension to the Lyne et al. [111] model in which the system switches between the two states four times per cycle, giving four independent switching periods. In this figure, $t_{\text{tot}} = t_A + t_B + t_C + t_D$. For a description of the three panels see Figure 2.7.

In Section 5.3.2 we will consider this switching model further in the context of PSR B1828-11 and develop a complete model for the spin-down rate and beam-width modulations.

2.3.3.3 A simple test of the switching hypothesis

PSR B1828-11, which has a distinct doubly peaked spin-down, is one of two pulsars for which the authors of Lyne et al. [111] applied their magnetospheric switching hypothesis too. Since that result, several other pulsars have been found with a similar double-peaked structure [134, 135]. In the previous section we introduced a basic empirical model to explain this double-peak in the context of magnetospheric switching. We will now describe a simple method to test this switching hypothesis which so far, to our knowledge, has not been attempted, but could be applied to current observational data.

If the double peaked spin-down rates of PSRs B1828-11, B0919+06, or any other pulsar is due to the model proposed by Perera et al. [134], then the maximum spin-down rate of the lower of the two peaks is a function of the time-averaging process and not the pulsar itself. Specifically, if the time-averaging was shorter than the shortest period there would be no second lower peak: both peaks would have the same value, $\dot{\nu}$. Therefore, if one had access to the raw data used to produce the time-averaged spin-down rate, one could

repeat the time-averaging process varying the time-averaging baseline which we denote by T . Then we could imagine measuring the maximum spin-down rate of the two peaks in the time-averaged spin-down rate and taking their ratio

$$\mathcal{R} = \frac{\dot{\nu}_{\max \text{ of lower peak}}}{\dot{\nu}_{\max \text{ of larger peak}}} \quad (2.16)$$

Then, if the Perera switching model is to be believed, \mathcal{R} should depend on T .

To demonstrate this, we can simulate the result numerically. We will do this by varying the time averaging baseline for the simulation in Figure 2.9 and calculating \mathcal{R} . The numerical results are plotted in Figure 2.11. This clearly shows that for $T \leq 50$ days,

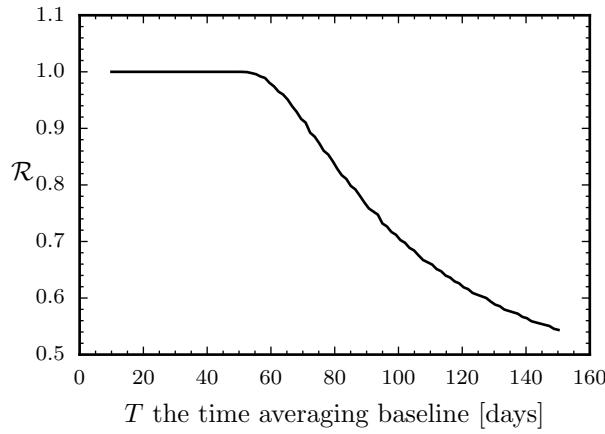


Figure 2.11: Demonstration of the variations in \mathcal{R} , defined in Eqn. (2.16), in response to changing the time averaging baseline, T .

when the time averaging baseline is shorter than the shortest switching duration t_C (see Figure 2.10), $\mathcal{R} = 1$ because the two peaks are of equal size. For $T > 50$ days, \mathcal{R} decreases until $T = 150$ days, above this the time-averaging is of a similar length to the total period of switching and so we can't resolve any of the features.

We have shown that variations in the ratio of the two peaks in spin-down rate would provide clear evidence in support of the Perera et al. [134] model. Of course in this example we have considered a simple situation where we have known a priori which peak is being underestimated due to the observer-method of calculating the time averaged spin-down rate. In real data, the situation may be more complicated: it could be that it is the trough which is overestimated instead of the peak being underestimated; or more than one switching duration may be shorter than the time averaging baseline. Nevertheless, the changing ratio of the two peaks (or troughs) may still prove a simple way to test the Perera model.

Appendix 2.A Toy model of stochastic resonance: particle in a potential

Here we present a simple toy model of stochastic resonance. This is a statistical phenomena occurring when a weak periodic forcing function is amplified by noise (see Jung [94] for a full treatment). The application to neutron stars and the problem of timing noise was first discussed by Cordes [45]. Here, we simply aim to describe the essential features of stochastic resonance (not its application to neutron stars).

We will consider a 1-dimensional space with a particle at a position x . The particle is subject to some potential and acted upon by a forcing function $F(t)$. In general, x could be any state variable of the system.

First consider the static case of a particle in a potential $U(x)$ given by:

$$U(x) = \frac{x^4}{4} - \frac{x^2}{2}. \quad (2.17)$$

This potential is characterised by two wells at $x = \pm 1$, a maximum exists between them at the origin. The particle in one of the wells sees a potential barrier ΔU corresponding to the height of the maximum above its position.

Assume the particle is acted upon by a random forcing function $F(t)$ which is modelled as a Gaussian white noise with standard-deviation σ . Depending on the magnitude of σ with respect to the potential, the motion of the particle admits two distinct cases:

1. $\sigma \ll \Delta U$ in which case the particle remains inside whichever well it initially starts in and does not escape.
2. $\sigma \gg \Delta U$ in this case the particle will not see the individual wells only the larger one.

The motion of the particle obeys the following equation of motion:

$$\frac{dx}{dt} = -\frac{\partial V(x, t)}{\partial x} + F(t). \quad (2.18)$$

The motion of the particle has two components, the deterministic effect of the potential and random fluctuations.

We now modify the potential to be acted on by a weak periodic function; this introduced a third possible type of behaviour. Writing the time dependent potential as

$$V(x, t) = \frac{x^4}{4} - \frac{x^2}{2} + \varepsilon x \cos(\omega_0 t). \quad (2.19)$$

Inserting this potential into the equations of motion:

$$\frac{dx}{dt} = x - x^3 + F(t) + \varepsilon \cos(\omega_0 t). \quad (2.20)$$

Solving this numerically we fix $\varepsilon = 0.001$, $\omega_0 = \frac{2\pi}{10}$ and choose three values of σ which illustrate typical behaviours of the solution. The first and last runs replicate the behaviour

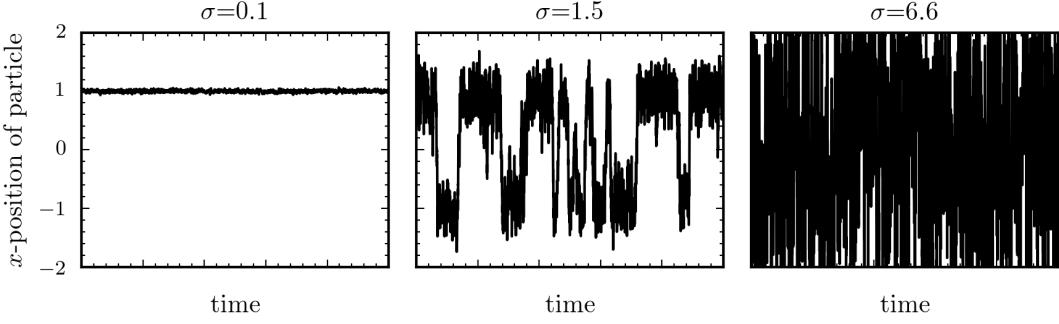


Figure 2.12: Three solutions to Eqn. (2.20) changing the random forcing functions strength σ . The left and right panels show the non-resonant solutions for the particle position: either the forcing function is weak compared to the potential and so the particle remains in well in which it begins; or the forcing function is much stronger than the potential and so the particle freely moves about the two wells. The middle panel illustrates the special case of stochastic resonance whereby the periodic fluctuations of the potential allow quasi-periodic variations in the particle's position between the two wells.

expected for a static well, either the particle is confined to the well it starts in, or the random noise is too strong and the individual wells are not observed. The middle case displays strong stochastic resonance: the solution displays a switching between bi-stable states but does not strictly follow the period of the forcing function. The important point here is that the forcing function may be weak, but provided it is periodic or at least quasi-periodic the signal is amplified by the random noise such that it may be visible in data sets where it would typically be considered lost.

Chapter 3

Action of the electromagnetic torque on a precessing neutron star

The rotation of an isolated neutron star is gradually slowed by torques. For most stars, the dominant torque is thought to be the electromagnetic torque due to an inclined magnetic dipole [132, 74]. In Section 2.3.2, we introduced free precession, the most general motion for a non-axisymmetric body. In later chapters we will be considering precession as a potential candidate to explain long term periodicities seen in pulsar timing residuals. Therefore, it is appropriate to first understand how precession manifests under the action of an electromagnetic torque and discuss the alignment of various axes of the star during its lifetime (as first investigated for a biaxial star by Goldreich [69]).

The full form of the electromagnetic torque was first written down by Deutsch [54] where it was found there were two distinct components: the spin-down component and the anomalous component (defined in Section 3.1). The subject of precession under the electromagnetic torque has already been studied, but many authors neglect the anomalous component in order to simplify the analytic calculations. In this chapter, we will build a numerical model solving the Euler rigid-body equations with the Deutsch torque. This will allow us to compare solutions in which the anomalous torque is ‘switched off’ with solutions to the full torque; in doing so we hope to better understand the importance of the anomalous torque and verify that in cases where it has been neglected (see for example [69]) that such an assumption was appropriate. A similar numerical model was studied by Melatos [122], although the authors did not switch on and off the anomalous component. Furthermore, the authors found solutions which they labelled as ‘persistent precession’: in Section 3.4 we will confirm that such solutions exist, but give a physical explanation from which we conclude that precession is not persistent.

The rest of this chapter is organised as follows: in Section 3.1 we will define the Euler rigid-body equations and the Deutsch [54] torque. In Section 3.2, we will discuss the simple case of a spherical star, then in Section 3.3 and Section 3.4 we will discuss a biaxial star with and without the anomalous part of the torque; in these sections we

will give an intuitive geometric way to understand the alignment predictions made by Goldreich [69]. Finally, we will conclude the chapter in Section 3.5.

3.1 Defining the model

Euler's rigid-body equations describe the motion of a rigid-body in a rotating reference frame. To derive these, we begin with Euler's second law for the angular momentum \mathbf{J} in the inertial frame

$$\left(\frac{d\mathbf{J}}{dt} \right)_{\text{in}} = \mathbf{T}, \quad (3.1)$$

where \mathbf{T} is the applied torque. The angular momentum is the product of the moment of inertia tensor and the spin-vector, $\mathbf{J} = I\boldsymbol{\Omega}$; in the inertial frame both of these quantities vary in time. We can simplify by transforming to a frame fixed in the rotating frame of the rigid-body and aligning the axes with the principal axes of I . In such a frame, the moment of inertia tensor is constant and diagonal. The transformation to a rotating reference frame requires a modification of the time derivative in Euler's second law (see for example §36 of Landau and Lifshitz [100] for a detailed discussion). The result is Euler's rigid-body equations in the rotating frame

$$\frac{d\mathbf{J}}{dt} + \boldsymbol{\Omega} \times \mathbf{J} = \mathbf{T}. \quad (3.2)$$

We will consider a biaxial body with the moment of inertia tensor defined in Eqn. (2.8) which has principal moments given by

$$I_{xx} = I_{yy} = I_0 \quad I_{zz} = I_0(1 + \epsilon_I). \quad (3.3)$$

For neutron stars, their extreme gravity ensures that $\epsilon_I \ll 1$. For a uniform density incompressible star, we take a typical moment of inertia to be

$$I_0 = \frac{2}{5}MR^2 \approx 10^{45} \text{g cm}^2 M_{1.4} R_6^2, \quad (3.4)$$

where $M_{1.4} = 1.4M_\odot$ and $R_6 = 10^6$ cm are typical values.

In Cartesian coordinates, Eqn. (3.2) gives us three coupled ODEs for Ω_x , Ω_y and Ω_z . In the torque-free case, analytic solutions can be found as we demonstrated in Section 2.3.2.

Neutron stars are acted on by torques from both their electromagnetic (EM) dipole and, if $\epsilon_I \neq 0$, they may also be acted on by torques from the emission of gravitational waves. For now, we will consider only the former effect which probably dominates for isolated neutron stars observed as pulsars. To model this EM torque, we imagine a dipole with magnitude m pointing along a unit vector \hat{m} frozen into the star. In the rotating frame, the body is axially symmetric about the \hat{z} axis; therefore we are free to fix the dipole in the $\hat{x} - \hat{z}$ plane at an angle χ to the \hat{z} axis without loss of generality. Often in the literature the effect of such an EM torque is modelled by a vacuum dipole torque when modelling neutron stars, but Deutsch [54] demonstrated that in a real finite volume (as opposed to a point-like) star, the dipole torque has two components: the usual spin-down

component and a second so-called anomalous component. We present the full torque here in the form found in Goldreich [69],

$$\mathbf{T} = \frac{2R}{3c} I_0 \epsilon_A \omega^2 (\boldsymbol{\Omega} \times \hat{\mathbf{m}}) \times \hat{\mathbf{m}} + \epsilon_A I_0 (\boldsymbol{\Omega} \cdot \hat{\mathbf{m}}) (\boldsymbol{\Omega} \times \hat{\mathbf{m}}), \quad (3.5)$$

where $\omega = |\boldsymbol{\Omega}|$ and ϵ_A is an effective magnetic ‘deformation’ as defined by Glampedakis and Jones [66]; it is of the order of the ratio of the magnetostatic energy to the rest-mass energy of the star and is given by

$$\epsilon_A = \frac{m^2}{I_0 R c^2} \approx 10^{-7} \left(\frac{B_0}{10^{15} \text{G}} \right)^2, \quad (3.6)$$

where in the second step we have expressed it in terms of the surface magnetic dipole field strength. Alternatively, it may be useful to solve for the surface magnetic field strength

$$B_0 = \left(\frac{4\epsilon_A I_0 c^2}{R^5} \right)^{1/2}. \quad (3.7)$$

The first term in the Deutsch torque is often referred to as the *spin-down*, or *braking* torque. As the name suggests, it is responsible for the power law retardation of spin frequency and has an associated timescale τ_S . The second term is known as the *anomalous* torque which acts on a timescale τ_A . We can then define three timescales (including the free precession timescale as derived in Eqn. (2.14)) which relate to this system, these are initially given by

$$\tau_P = \frac{2\pi}{\epsilon_I \omega_0 \cos(a_0)}, \quad (3.8)$$

$$\tau_A = \frac{2\pi}{\epsilon_A \omega_0}, \quad (3.9)$$

$$\tau_S = \frac{3c}{2R} \frac{1}{\epsilon_A \omega_0^2}. \quad (3.10)$$

The spin-down age defined here, τ_S , is equivalent to the original definition of τ_{age} in Eqn. (1.10) up to a factor of $2/\sin^2(\alpha)$ where α is the angle between the spin-vector and the magnetic dipole. Since we are considering cases where this angle evolves, we will parameterise by τ_S in this chapter.

The ordering of these timescales characterises the type of solutions found to Eqn. (3.2) with the torque of Eqn. (3.5). There are six potential orderings, however by considering the ratio $\tau_A/\tau_S \sim v/c$, where v is the equatorial velocity, we see that $\tau_A \ll \tau_S$ reducing the number of possible orderings to three.

We now have the basic components of a neutron star: a biaxial rigid-body spun down by a dipole torque. To help familiarise the reader with the notation, in Table 3.1 we list the *spherical coordinates* (the magnitude, polar angle, and azimuth) of the spin-vector and magnetic dipole and also show their projection onto the $\hat{\mathbf{x}} - \hat{\mathbf{z}}$ plane in Figure 3.1. We work in a system of coordinates where the polar angle is that made with the $\hat{\mathbf{z}}$ axis

and the azimuthal angle is the angle between the projection of the vector onto the $\hat{x} - \hat{y}$ plane and the \hat{x} axis.

		Magnitude	Polar angle	Azimuthal angle
Spin-vector	Ω	ω	a	φ
Magnetic dipole	m	m	χ	0

Table 3.1: Spherical coordinates of the spin-vector and magnetic dipole.

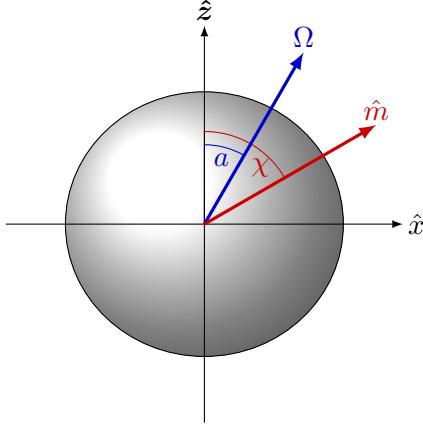


Figure 3.1: A sketch of the $\hat{x} - \hat{z}$ plane, the magnetic dipole lies solely in this plane such that its azimuthal component is always zero. Only the projection into this plane of the spin-vector is shown. In general it has an azimuthal component given by φ .

In the following sections we will develop our intuition for the different types of solutions to this model. These results are calculated by exact numerical evolution of Eqn. (3.2) with the torque defined in Eqn. (3.5). The simulations model a neutron star with a magnetic field of $B_0 \approx 10^{13}$ G and a radius $R = 10^6$ cm. We must also provide an initial condition for the angular frequency ω . This is not the spin frequency at the birth of the neutron star, but the frequency at the start of the observation; typical values are of the order ~ 10 rad/s. However, at this frequency the numerical solutions are computationally expensive. To reduce the time required to produce a solution we will work with an initial frequency of $\omega = 10^4$ rad/s. This does not change the form of solution, only the timescales over which they evolve.

The rest of this chapter is organised as follows: in Section 3.2 we will begin by discussing a spherical star. Since we particularly aim to understand the role of the anomalous torque, in Chapter 3.3 we will investigate a biaxial star acted on by only the spin-down component of the torque. In Chapter 3.4 we then repeat the investigation of a biaxial star, but this time include the anomalous torque and compare the results. In doing this we hope to better understand the role of the anomalous torque on precession.

3.2 Spherical star

We begin by considering the simplest case: a spherical star with $\epsilon_I = 0$ under the influence of the spin-down torque only. This model was first considered by Davis and Goldstein [52] and separately by Michel and Goldwire [125]; both sets of authors discovered that torque causes the spin axis to align with the magnetic axis on the spin-down timescale.

To observe this behaviour, we set $\chi = 30^\circ$ with the spin-vector initially at $a = 50^\circ$ and $\epsilon_A = 5 \times 10^{-11}$. Evolving the model numerically, we plot the spherical coordinates of the spin-vector in Figure 3.2. This shows a brief spin-down which halts once the spin-vector aligns with the magnetic dipole $a \rightarrow \chi = 30^\circ$. The azimuthal angle φ does not vary; this is expected since precession, a monotonic increase in φ , does not occur for spherical bodies. This result agrees with the analytic calculations: the alignment occurs on a few times the spin-down time scale $\tau_S = 9 \times 10^6$ s.

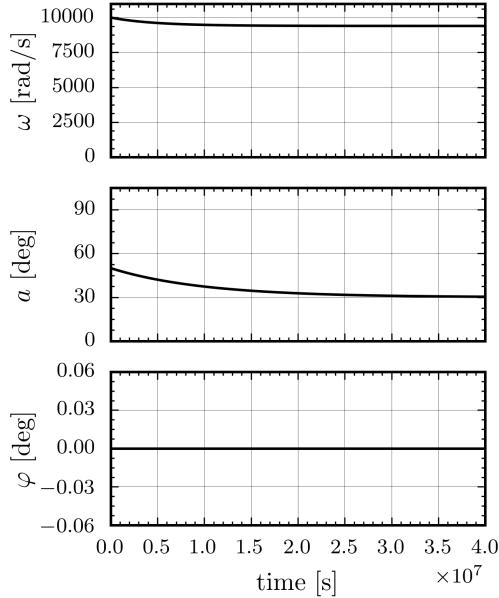


Figure 3.2: Plot of the spherical coordinates of the spin-vector for a spherical star. Note that the magnetic dipole is inclined at $\chi = 30^\circ$ to the \hat{z} axis.

3.3 Biaxial neutron star with no anomalous torque

We will now consider a biaxial star with $\epsilon_I \neq 0$. But to first build some understanding we will remove the anomalous part of the torque. This is done so that later on when we consider solutions in the presence of the anomalous torque, we understand which effects are due to the anomalous part and which to the spin-down part. Neglecting the anomalous torque reduces the number of timescales to just two; the spin-down and the precession timescale. The solutions are categorised by the ordering of these timescales;

it is shown in appendix 3.A that these orderings remain valid for the duration of the spin-down. The ordering of the two timescales gives three different types of solutions corresponding to three cases:

$$\text{Regions A: } \tau_S > \tau_P \quad \text{Region B: } \tau_S \sim \tau_P \quad \text{Region C: } \tau_S < \tau_P \quad (3.11)$$

Since all solutions in each region display the same characteristic behaviour, we will illustrate this by picking three ‘neutron stars’ and presenting results for each of these. Each of these pulsars, defined by the parameters listed in Table 3.2 will help us to understand the types of solutions for each ordering of the timescales.

Star	ϵ_I	ϵ_A	$B_0[\text{Gauss}]$	$\tau_S [\text{s}]$	$\tau_P [\text{s}]$
A	1.0×10^{-9}	5.0×10^{-11}	1.3×10^{13}	9.0×10^6	9.8×10^5
B	4.0×10^{-11}	5.0×10^{-11}	1.3×10^{13}	9.0×10^6	2.4×10^7
C	1.0×10^{-15}	5.0×10^{-11}	1.3×10^{13}	9.0×10^6	9.8×10^{11}

Table 3.2: Table of ϵ_I and ϵ_A and the inferred magnetic field and initial timescales for the three simulated stars. In calculating the magnetic fields and timescales, we have used a canonical value of $R = 10^6$ cm for the radius, the initial value of the spin frequency, $\omega_0 = 10^4$ rads/s and the initial value of $a_0 = 50^\circ$. As ω and a evolve, the timescales of each star will also evolve from these initial values.

The phase space of solutions is plotted in Figure 3.3, parameterised by the elastic deformation ϵ_I and the surface magnetic field

$$B_0 = \frac{2m}{R^3}. \quad (3.12)$$

This illustrates how the phase space is sliced into separate regions by the ordering of the timescales.

Without the anomalous torque, the ordering of the timescales defines two distinct regions separated by $\tau_S = \tau_P$. In using an abnormally large value for the initial angular spin frequency ($\omega_0 = 10^4$ [rads/s]), we have skewed this phase space - that is, it cannot be compared with the real pulsar population. Using a more realistic value, the character of solutions remains the same - only the timescales vary.

The elastic deformation ϵ_I may take either a positive or a negative value; these correspond to either an oblate or prolate mass distribution. In the following work, we use only positive values of ϵ_I . It will be noted when this is important, but otherwise solutions are invariant to the sign of ϵ_I . Finally, we must choose the angle between the magnetic dipole and the elastic deformations χ . We choose two values (30° and 75°) to demonstrate important cases discussed in the next section. Initial conditions are thought to have some bearing on the solutions [see 122]. In the current work, we will start with the spin-vector having an inclination angle of $a_0 = 50^\circ$ and azimuth of $\varphi = 0^\circ$.

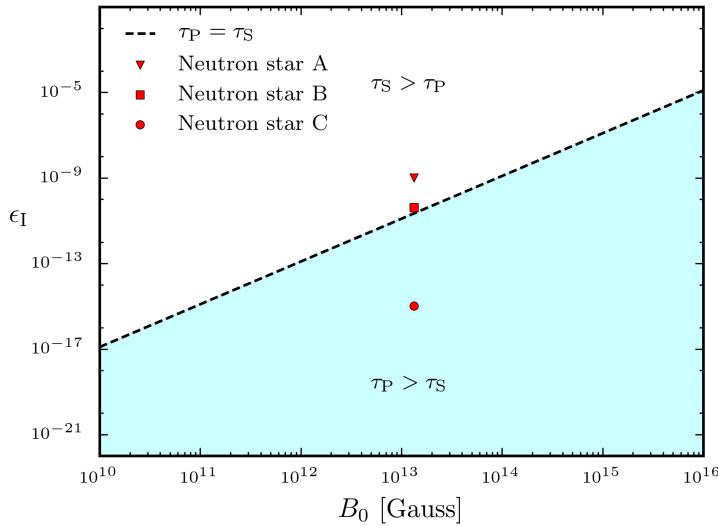


Figure 3.3: Phase space of solutions categorised by the ordering of the two timescales as a function of the elastic deformation ϵ_I and the surface magnetic fields B_0 .

3.3.1 Neutron star A in region $\tau_S \gg \tau_P$

Neutron stars in this region are described by the work of Goldreich [69]. In addressing the shortcomings of the analytic spherical solutions [52, 125], Goldreich considered a neutron star with a solid crust capable of supporting elastic strains. In order to find an analytic solution, Goldreich assumed that the precession timescale was significantly shorter than the spin-down timescale $\tau_S \gg \tau_P$. Averaging over equations (3.2) with this assumption yields

$$\left\langle \frac{d\omega}{dt} \right\rangle = -\frac{2R}{3c} \epsilon_a \omega^3 \left[\sin^2 \chi + \sin^2 a \left(1 - \frac{3}{2} \sin^2 \chi \right) \right], \quad (3.13)$$

$$\left\langle \frac{da}{dt} \right\rangle = -\frac{2R}{3c} \epsilon_a \omega^2 \sin a \cos a \left(1 - \frac{3}{2} \sin^2 \chi \right). \quad (3.14)$$

Goldreich realised that the second equation tells us whether the polar angle a will either grow or decay on the spin-down timescale depending on whether χ was greater or less than a critical value $\chi_{\text{cr}} \approx 55^\circ$.

We first present results for all the spherical coordinates during the spin-down in Figure 3.4. For the angle χ , we use two values, 30° and 75° , and find that a tends to either 0 or $\pi/2$ on the spin-down timescale, in agreement with Goldreich [69]. During this alignment, the spin magnitude decays and φ monotonically increases corresponding to the spin-vector precessing about the $\hat{\mathbf{z}}$ axis.

In Figure 3.5, we investigate the differences between solutions to Goldreich's averaged equations (3.14) and the exact solution without averaging. This shows that both agree on the long term behaviour occurring on the spin-down timescale. However, in the exact solutions we see additional oscillations at the faster precession timescale; this precession

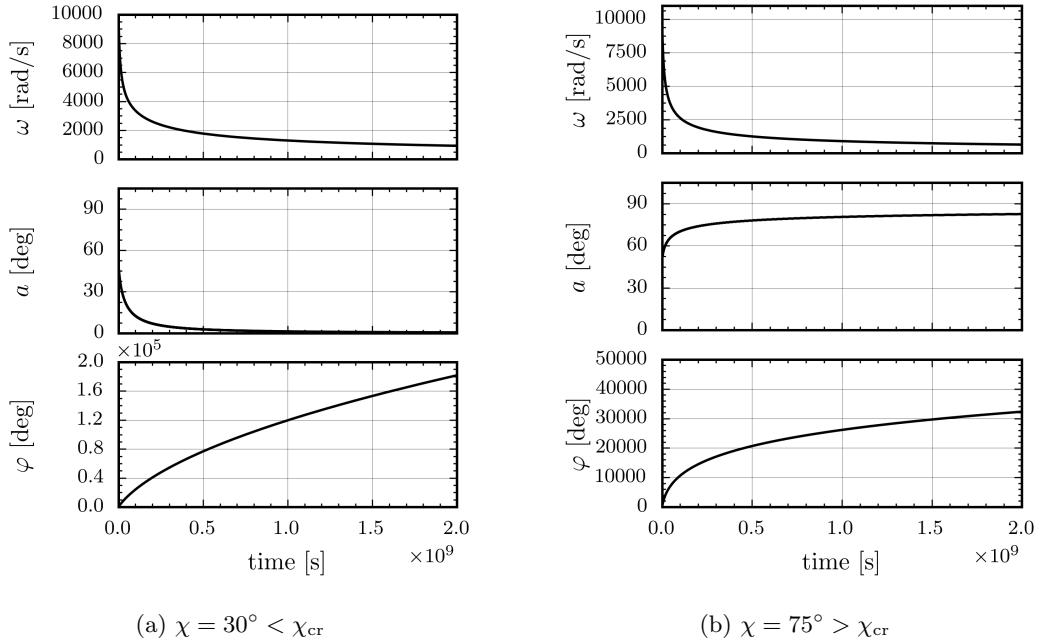


Figure 3.4: Plot of the spherical coordinates of the spin-vector for neutron star A without the anomalous torque component. The two choices of χ allow us to confirm Goldreich's dependence of the alignment of the spin-vector on χ : eventually we see that $a \rightarrow 0$ for $\chi = 30^\circ$ while $a \rightarrow 90^\circ$ for $\chi = 75^\circ$.

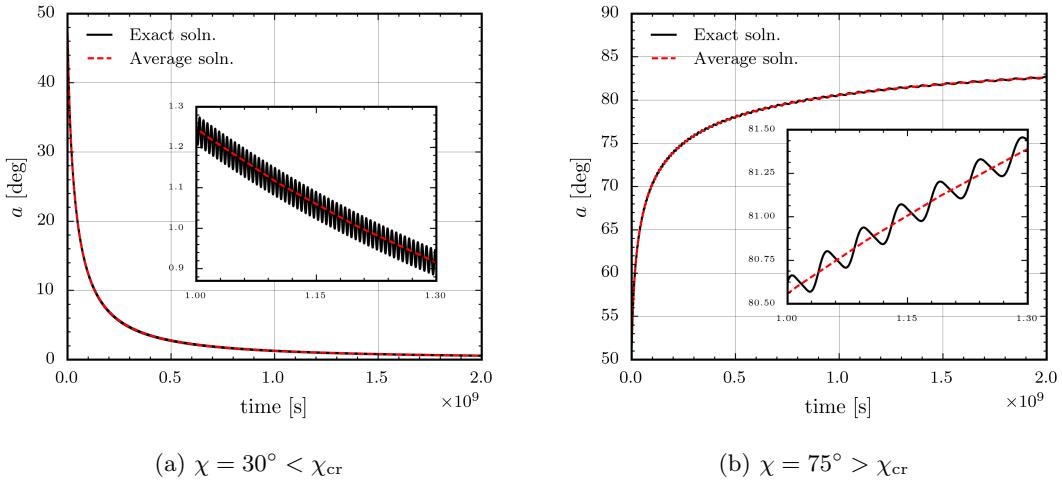


Figure 3.5: Plot of the angle a comparing the numerical solution of Eqn. (3.2) and the solution to the averaged equation given in Eqn. (3.14) for neutron star A without the anomalous torque component.

timescale is longer than the initial value defined in Table 3.2 due to the spin-down of the star. Moreover, since the two values of χ produce opposing evolutions in a , the precession timescale is also different in the two solutions being shorter for $\chi = 30^\circ$ than for $\chi = 75^\circ$. In contrast, the averaged equations do not include these short term oscillations because Goldreich averaged over the precessional modulations.

A geometric interpretation of the critical value χ_{cr}

The significance of the critical value $\chi_{\text{cr}} \sim 55^\circ$ has not been discussed in the literature. We will now develop a novel geometric argument, based on the intuitive ideas in §37 of Landau and Lifshitz [100], which provides a way to interpret this value as a natural consequence of the conservation of energy and angular momentum.

For a biaxial mass distribution as described above, over short timescales compared to the spin-down, the conservation of energy requires that

$$1 = \frac{J_x^2}{2I_{xx}E} + \frac{J_y^2}{2I_{yy}E} + \frac{J_z^2}{2I_{zz}E} = \frac{J_x^2}{2I_0E} + \frac{J_y^2}{2I_0E} + \frac{J_z^2}{2I_0(1+\epsilon_I)E}, \quad (3.15)$$

where J_i are the components of the angular momentum. For the conservation of momentum, we have that

$$J^2 = J_x^2 + J_y^2 + J_z^2. \quad (3.16)$$

Both of these conservation equations have a geometric interpretation in the Cartesian momentum space: Eqn. (3.15) describes a biaxial ellipsoid with semi-axes given by $\sqrt{2I_0E}$, $\sqrt{2I_0E}$ and $\sqrt{2I_0(1+\epsilon_I)E}$ centred on the origin, while Eqn. (3.16) describes a sphere of radius J also centred on the origin. The ellipsoid is *oblate* if $\epsilon_I > 0$ while it is *prolate* if $\epsilon_I < 0$. To keep the following discussion simple, we will assume that $\epsilon_I > 0$ and then we will discuss this choice at the end of this section.

Physical solutions require the conservation of both the energy and momentum and therefore possible solutions lie on the intersection of the sphere and ellipsoid. In order that the two intersect and hence both conservation laws are satisfied, the radius of the sphere must be bound by

$$2EI_0 < J^2 < 2EI_0(1+\epsilon_I). \quad (3.17)$$

If the energy and momentum of the system are fixed, the intersection of the sphere (describing the conservation of momentum) and ellipsoid (describing the conservation of energy) always describes a circle about the \hat{z} axis. To illustrate this, in Figure 3.6 we plot the intersection for three different sizes of the sphere holding the ellipsoid fixed. The intersection describes exactly the solutions which satisfy both the conservation equations and therefore the set of all physical solutions.

Over times much shorter than the spin-down timescale, the energy and momentum of our biaxial star acted on by the spin-down torque are fixed. As we have seen above, in order to satisfy the conservation equations this means that the angular momentum vector J_i has a class of possible solution consisting of a ring about the momentum-space \hat{z} . Since $J_i = I_{ij}\omega^j$, the star is biaxial, and moment of inertia tensor is static. Solutions for the spin-vector ω^j must also exist as points on a related circle. This related circle is exactly the precession of ω about the \hat{z} axis.

Both the energy and the momentum decay in time due to the spin-down. As a result both the sphere and ellipsoid will shrink, but not necessarily at the same rate. If we imagine observing the ellipsoid such that it appears to be fixed, we may see the sphere

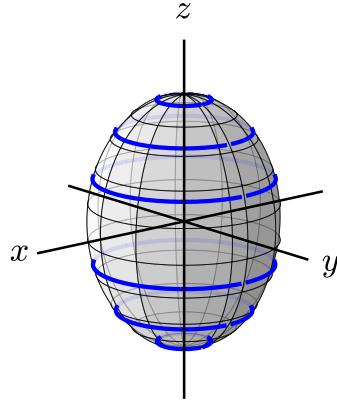


Figure 3.6: Intersections of the ellipsoid and sphere defined in Eqn. (3.15) and Eqn. (3.16) respectively for three different values of M at fixed E . The spheres of radius M are not visible, only the blue line shows where they would intersect the fixed ellipsoid.

either shrink or grow with respect to the ellipsoid. We can parameterise the relative rate of shrinking by considering the quantity

$$A(t) = \frac{J^2}{2EI_0}. \quad (3.18)$$

We could have chosen any one of the ellipsoid's semi-axes since they are all proportional up to factors of ϵ_I . The evolution of the spin-vector is related to the rate of change of A . This quantity determines whether the sphere shrinks or grows with respect to the ellipsoid. It is worth taking a moment to understand the different cases which are parameterised by the sign of \dot{A} :

1. $\dot{A} > 0$ The sphere grows with respect to the ellipsoid. In this case, the intersection circles will ‘close up’ around the \hat{z} axis. This agrees with the solution for $\chi = 30^\circ$ given in Figure 3.5(a). The angle between ω and the \hat{z} axis tends to zero while φ monotonically increases. This describes the spin-vectors following a circle about the \hat{z} axis, which gradually decreases in radius.
2. $\dot{A} < 0$ The sphere shrinks with respect to the ellipsoid. In this case the intersection circles will increase in radius until $J \rightarrow \sqrt{2EI_0}$. This describes exactly the solution in Figure 3.5(b): that is $a \rightarrow 90^\circ$ while φ monotonically increases.

The sign of \dot{A} tells us about the long-term evolution of the system. To determine it, we begin by differentiating Eqn. (3.18)

$$\dot{A} = \frac{1}{(2EI_0)^2} \left(\frac{d}{dt}(J^2)2EI_0 - \frac{d}{dt}(2EI_0)J^2 \right), \quad (3.19)$$

simplifying and writing in index notation such that $J^2 = J^i J_i$, we have that

$$\dot{A} = \frac{1}{2E^2 I_0} (2J^i \dot{J}_i E - \dot{E} J^i J_i). \quad (3.20)$$

Now the energy is purely kinetic such that $E = \frac{1}{2} I_{ij} \omega^i \omega^j$, and therefore $\dot{E} = T^i \omega_i$. Then, noting that $\dot{J}^i = T^i$, we have

$$\dot{A} = \frac{1}{E^2} (J^i T_i I_{jk} \omega^j \omega^k - T^i \omega_i J^j J_j). \quad (3.21)$$

For a diagonalised matrix, I_{ij} has only 3 non zero components and therefore $I_{jk} \omega^j \omega^k = (I\omega)_j \omega^j$. In addition, we can write $J^i = (I\omega)^i$ and so we find

$$\dot{A} = \frac{1}{E^2} T^i ((I\omega)_i \omega_j - \omega_i (I\omega)_j) (I\omega)^j, \quad (3.22)$$

finally simplifying to

$$\dot{A} = \frac{1}{E^2} T^i \omega_i \omega_j (I\omega)^j (I_i^i - I_j^j). \quad (3.23)$$

Expanding the summation on the right hand side, all terms for which $i = j$ will be zero. Working with a biaxial mass distribution where $I_1 = I_2$, then the only non-zero terms are given by

$$\begin{aligned} T^i \omega_i \omega_j (I\omega)^j (I_i^i - I_j^j) &= T_1 \omega_1 I_0 (1 - \epsilon_1) \omega_3^2 (I_0 - I_0 (1 + \epsilon_1)) \\ &\quad + T_2 \omega_2 I_0 (1 - \epsilon_1) \omega_3^2 (I_0 - I_0 (1 + \epsilon_1)) \\ &\quad + T_3 \omega_3 (I_0 (1 + \epsilon_1) - I_0) (I_0 \omega_1^2 + I_0 \omega_2^2), \end{aligned} \quad (3.24)$$

$$= I_0^2 \epsilon_1 \omega_3 (T_3 (\omega_1^2 + \omega_2^2) - \omega_3 (1 + \epsilon_1) (T_1 \omega_1 + T_2 \omega_2)). \quad (3.25)$$

Working up to 1st order in ϵ_1 and transforming to spherical polar coordinates, we then insert the spin-down torque and neglect the anomalous torque. This yields

$$\begin{aligned} \dot{A} E^2 &= I_0^2 \epsilon_1 \omega^3 \left(\frac{2R}{3c} \epsilon_A \omega^3 \right) \cos a ([\sin \chi \cos \chi \sin a \cos \varphi - \sin^2 \chi \cos a] \sin^2 a \\ &\quad - \cos a ([\sin \chi \cos \chi \sin a \cos \varphi - \sin^2 \chi \cos a] \cos \varphi \sin a - \sin^2 a \sin^2 \varphi)). \end{aligned} \quad (3.26)$$

Now we work in the limit $\tau_P \ll \tau_S$ where φ varies on the precession timescale, which is much faster than the evolution in either a or ω . Therefore, we can average over a free precession period and drop the constant coefficients:

$$\langle \dot{A} \rangle \propto \cos a \left(-\sin^2 a \cos a \sin^2 \chi + \frac{1}{2} (\sin^2 a \cos a \cos^2 \chi + \sin^2 a \cos a) \right) \quad (3.27)$$

$$\propto \cos^2 a \sin^2 a \left(1 - \frac{3}{2} \sin^2 \chi \right). \quad (3.28)$$

Finally, putting all the coefficients back in, we have an expression for the averaged rate of change of A under the same assumptions made by Goldreich [69]

$$\langle \dot{A} \rangle = \frac{1}{\langle E \rangle^2} I_0^2 \epsilon_1 \omega^6 \frac{2R}{3c} \epsilon_A \cos^2 a \sin^2 a \left(1 - \frac{3}{2} \sin^2 \chi \right). \quad (3.29)$$

This result suggests that only two factors determine the sign of $\langle \dot{A} \rangle$: the sign of ϵ_I and the term $1 - \frac{3}{2} \sin^2 \chi$. However, for the argument which led to this conclusion, we assumed that $\epsilon_I > 0$. Repeating the argument with $\epsilon_I < 0$ (starting from Eqn. (3.17) which would need to be rearranged to reflect this change), we find that Eqn. (3.29) is multiplied by an overall minus sign. Therefore, the only factor which determines the sign of $\langle \dot{A} \rangle$ is in fact $1 - \frac{3}{2} \sin^2 \chi$, from which we find the critical value $\chi_{\text{cr}} \approx 55^\circ$; this is in agreement with the findings of Goldreich [69]. Moreover, by calculating the condition in this way, we have found a new interpretation for the critical value as a tipping point between the rates of change of energy and angular momentum.

3.3.2 Neutron star B in region $\tau_S \sim \tau_P$

Unlike the previous examples, there is no comparison in the literature which considers $\tau_P \sim \tau_S$ while neglecting the anomalous torque. This is primarily because the similarity of the timescales means simplifying assumptions cannot be made. The results have not been included here as they are exactly alike with the results of neutron star A, except that the precession, the alignment of the spin-vector, and the spin-down all happen on the same timescale.

3.3.3 Neutron star C in region $\tau_S \ll \tau_P$

In neutron star C, the timescales take the ordering $\tau_P \gg \tau_S$. The limit of this region is the spherical star ($\epsilon_I \rightarrow 0$) discussed in section 3.2; as such we should expect the results to be similar. Solving the equations of motion for neutron star C, the spherical coordinates of the spin axis are given in Figure 3.7. Both choices of χ yield a similar result; the polar angle a tends to χ on the spin-down timescale as anticipated by the likeness to the spherical star. In contrast, the azimuthal angle φ appears to begin increasing monotonically before decreasing to a constant value. It is thought that this is due to the small but non-vanishing magnitude of ϵ_I . In the perfectly spherical limit, we observe no such increase in φ . Therefore, the small deformation offsets the steady state solution of the spin axis from the magnetic dipole.

For an intuitive understanding consider that the torque acts to align the spin-vector with the magnetic dipole; when this occurs the torque vanishes. Since the mass is not spherical, non-precessing solutions exist when the spin-vector aligns with the principal axes of the moment of inertia tensor \hat{x}, \hat{y} and \hat{z} . So when both effects act, the steady state solution is an intermediary point between the two, weighted by the relative strength of each effect. In this case, the torque dominates and so the spin-vector lies close to, but not exactly aligned with, the magnetic dipole. During this time, the star spins down but upon alignment the spin frequency asymptotically approaches a constant non-zero value.

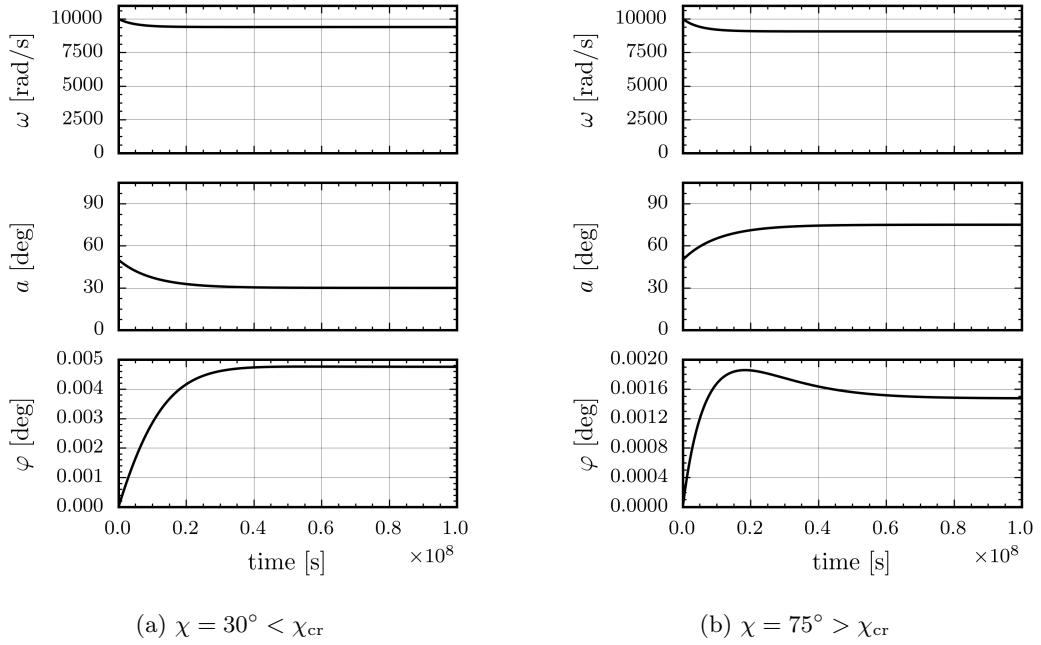


Figure 3.7: Plot of the spherical coordinates of the spin-vector for neutron star C without the anomalous torque component.

3.4 Biaxial neutron star including the anomalous torque

In the previous section we considered solutions to Eqn. (3.2) under the action of the torque in Eqn. (3.5) without the anomalous contribution. In this section, we will include this contribution and understand the effect it has on the character of solutions. However, before doing so, we will first consider the appropriate reference frame in which to view the system.

3.4.1 The effective body frame

Working with a biaxial body and the Deutsch torque, [122] discovered non-trivial solutions where the spin axis aligns with a non-zero angle between the magnetic dipole and the principal axis of the moment of inertia tensor. This was interpreted by Andrew Melatos as evidence for persistent precession. We show in Section 3.4.4 that in fact the spin-vector is aligning with another rotating reference frame that is held at a fixed angle to the original rotating reference frame. This reference frame, which we will call the *effective* body frame, is a direct result of the anomalous torque contribution in Eqn. (3.5); we will now show how this comes about from a rearrangement of the equations of motion as first noted by Glampedakis and Jones [66].

Writing Eqn. (3.2) with the Deutsch torque, we then manipulate it to give

$$J_{,t}^i + \epsilon^{ijk}\Omega_j J_k = N_{\text{spin-down}}^i + \epsilon_A I_0 \Omega^a \hat{m}_a \epsilon^{ijk} \Omega_j \hat{m}_k, \quad (3.30)$$

$$\Rightarrow J_{,t}^i + \epsilon^{ijk}\Omega_j (J_k - \epsilon_A I_0 \Omega^a \hat{m}_a \hat{m}_k) = N_{\text{spin-down}}^i, \quad (3.31)$$

$$\Rightarrow J_{,t}^i + \epsilon^{ijk}\Omega_j (I_{ka} - \epsilon_A I_0 \hat{m}_a \hat{m}_k) \Omega^a = N_{\text{spin-down}}^i. \quad (3.32)$$

By arranging the equation this way and making use of $J_i = I_{ij}\Omega^j$, we can write an *effective moment of inertia tensor* given by

$$I'_{jk} = I_{jk} - I_0 \epsilon_A \hat{m}_j \hat{m}_k, \quad (3.33)$$

$$= \begin{bmatrix} I_0(1 - \epsilon_A \sin^2 \chi) & 0 & -I_0 \epsilon_A \sin \chi \cos \chi \\ 0 & I_0 & 0 \\ -I_0 \epsilon_A \sin \chi \cos \chi & 0 & I_0(1 + \epsilon_I - \epsilon_A \cos^2 \chi) \end{bmatrix}. \quad (3.34)$$

Note that we will use the prime to denote the coordinates in the effective body frame.

We can calculate the eigenvalues of the effective moment of inertia tensor, which are given by

$$\lambda_2 = I_0, \quad \lambda_{\pm} = \frac{I_0}{2} \left(2 + \epsilon_I - \epsilon_A \pm \sqrt{\epsilon_A^2 + \epsilon_I^2 - 2\epsilon_A \epsilon_I \cos(2\chi)} \right). \quad (3.35)$$

If we diagonalise this effective moment of inertia tensor, these eigenvalues are the diagonal entries, and the associated eigenvectors are the principal axes of the effective body frame. The eigenvalues always take the ordering

$$\lambda_+ > I_0 > \lambda_-. \quad (3.36)$$

Defining the axes of the effective body frame by \mathbf{e}_i , it is natural to associate \mathbf{e}_2 with the $\hat{\mathbf{y}}$ of the body frame axes such that they are always parallel. Since the other axes must be orthonormal, the transformation must consist of a rotation in the $x - z$ plane by an angle β . We are free to set the \mathbf{e}_3 axis to always take the largest eigenvalue and then define β as the angle made by \mathbf{e}_3 with the $\hat{\mathbf{z}}$ axis. β is then given by

$$\beta = \arctan \left(\frac{\mathbf{e}_{3,1}}{\mathbf{e}_{3,3}} \right) = \arctan \left(\frac{\epsilon_I - \epsilon_A \cos(2\chi) - \sqrt{\epsilon_A^2 + \epsilon_I^2 - 2\epsilon_A \epsilon_I \cos(2\chi)}}{2\epsilon_A \sin \chi \cos \chi} \right). \quad (3.37)$$

In Figure 3.8, we provide a schematic of the various angles and relating the effective body frame to the body frame axes.

Understanding the effective body frame The effective body frame is a simple rotation of the axis defined by the moment of inertia tensor to accommodate the effects of the anomalous torque. To understand when this becomes significant, we plot β as a function of $|\epsilon_A/\epsilon_I|$ in Figure 3.9. When $|\epsilon_A \ll \epsilon_I|$, such that the anomalous torque effects are negligible, we recover the usual moment of inertia tensor. That is, $\beta = 0$ or $\beta = \pi/2$, dependent on the sign of ϵ_I . This is a consequence of choosing \mathbf{e}_3 to take the

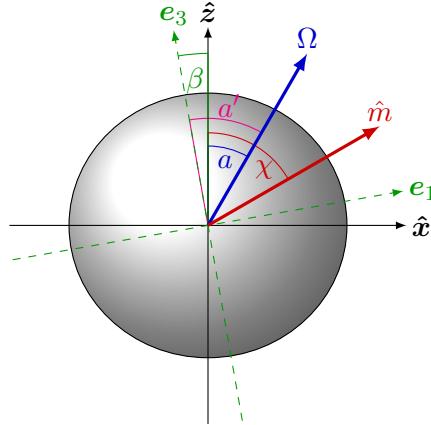


Figure 3.8: Schematic of the two sets of axes for an arbitrary β which in this is negative. Note that β is defined to be positive for a right hand rotation about the \hat{y} axis which is defined to be into the page.

largest eigenvalue. In the opposing limit $|\epsilon_A \gg \epsilon_I|$, in which the magnetic deformation dominates, the sign of the elastic deformation no longer splits the solutions and in both cases a tends to $\chi - 90^\circ$. This means the effective body frame always aligns the e_1 axis (which has the smallest eigenvalue) with the magnetic dipole.

We will show in the following sections that, when the anomalous torque is included, it is the effective body frame to which the system aligns on the spin-down timescale (much as it is the body frame to which the system aligns when the anomalous torque contribution was neglected in Section 3.3).

For the three neutron stars (A, B, and C as first used in Section 3.3) representative of different orderings of the timescales, all three have a non-zero β values. To familiarise the reader with the magnitudes in the different regions, in Table 3.3 we list the relevant timescales and β for each of the three neutron stars.

Star	ϵ_I	ϵ_A	B_0 [Gauss]	τ_S [s]	τ_A [s]	τ_P [s]	$\beta(\chi = 30^\circ)$	$\beta(\chi = 75^\circ)$
A	1.0×10^{-9}	5.0×10^{-11}	1.3×10^{13}	9.0×10^6	1.3×10^7	9.8×10^5	-1.3°	-0.69°
B	4.0×10^{-11}	5.0×10^{-11}	1.3×10^{13}	9.0×10^6	1.3×10^7	2.4×10^7	-35.0°	-8.4°
C	1.0×10^{-15}	5.0×10^{-11}	1.3×10^{13}	9.0×10^6	1.3×10^7	9.8×10^{11}	-60.0°	-15.0°

Table 3.3: Table of relevant values for the selected points updated from Table 3.2 to include the value of β , as defined in Eqn. (3.37).

In the following, we will present all results in the effective body frame, rather than the usual rotating frame. The solutions are first calculated by numerical evolution of the Euler rigid-body equations with the full torque Eqn. (3.5) in the usual rotating frame. Then, in the post-processing stage, we apply the effective body frame transformation (a rotation by the angle β).

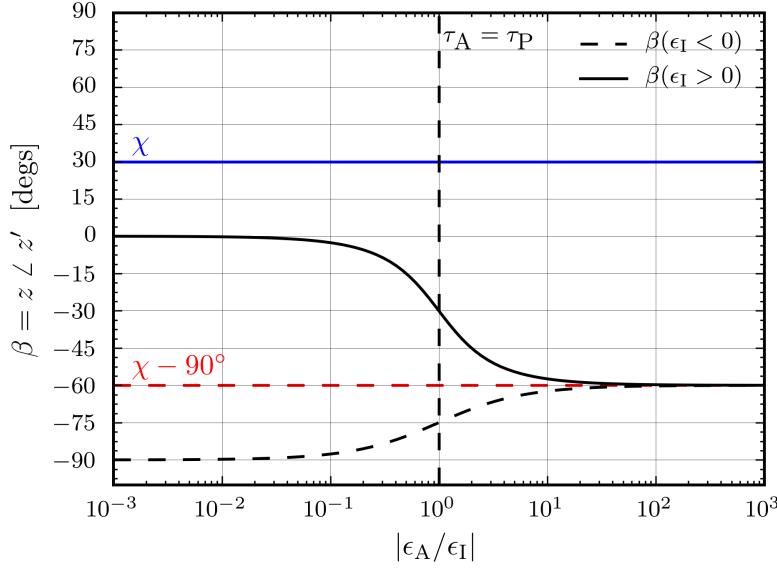


Figure 3.9: Plot of β as a function of the ratio $|\epsilon_A/\epsilon_I|$ for a prolate, $\epsilon_I < 0$, and oblate, $\epsilon_I > 0$, mass distribution with $\chi = 30^\circ$. In the limit $\epsilon_A \gg \epsilon_I$ for both the oblate and prolate cases, the effective body frame aligns with the magnetic dipole. In the opposing limit the magnetic dipole has little effect and the effective body frame aligns with the principal axis of the moment of inertia. There is a smooth transition between these two extremes when $\epsilon_A \sim \epsilon_I$. The splitting into oblate and prolate solutions results from choosing to associate e_3 with the largest eigenvalue.

3.4.2 Phase space

The introduction of the anomalous torque means we must also consider the anomalous torque timescale as given in Eqn. (3.9). Three timescales can take six orderings. However, the spin-down and anomalous timescale obey the following relation

$$\tau_S = \frac{3c}{2R} \frac{2\pi}{\omega_0} \tau_A. \quad (3.38)$$

The coefficient $3c/2R\omega_0$ effectively measures the importance of special relativity at the surface of the star. Since we expect $c \gg R\omega_0$, this means that only three physical orderings exist:

1. Region A: $\tau_S > \tau_A > \tau_P$
2. Region B: $\tau_S > \tau_P > \tau_A$
3. Region C: $\tau_P > \tau_S > \tau_A$

We can therefore update the phase space of possible solutions with the additional anomalous timescale; the result is given in Figure 3.10.

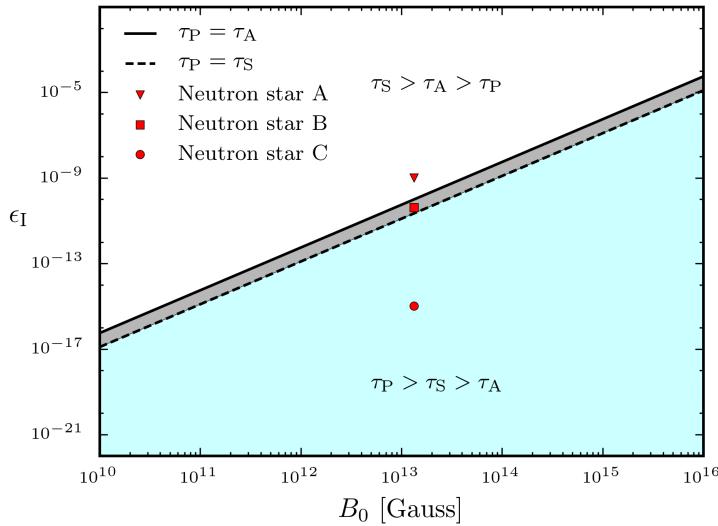


Figure 3.10: Phase space diagram including the anomalous torque contributions. This differs from Figure 3.3 in that there is an additional grey band marking ‘region B’ where $\tau_S > \tau_P > \tau_A$.

3.4.3 Neutron star A in region $\tau_S > \tau_A > \tau_P$

We expect the alignment of the spin axis to agree with the findings in Section 3.3.1 and study the changes in behaviour due to the introduction of the anomalous torque. Working in the effective body frame rotates the solutions (for star A) by a few degrees as shown in Table 3.3. Plotted in Figure 3.11 are the spherical coordinates of the spin-vector in the effective body. This can be compared with the result in Figure 3.4, which is the same system without the anomalous torque component.

The final state alignment agrees well with results neglecting the anomalous torque for both values of χ . For $\chi = 30^\circ$, the intermediate behaviour appears similar to those shown in Section 3.3.1. For $\chi = 75^\circ$, we have a large discontinuity in the solution suggesting a significant difference between solutions when including the anomalous torque. To understand the cause of this, we return to the geometric argument made in Section 3.3.1.

A geometric interpretation Working in the effective body frame for which the effective moment of inertia tensor is triaxial with principal moments of inertia given by Eqn. (3.35), the conservation of energy equation is

$$1 = \frac{J_1^2}{2\lambda_- E} + \frac{J_2^2}{2I_0 E} + \frac{J_3^2}{2\lambda_+ E}, \quad (3.39)$$

while the conservation of momentum remains as in Eqn. (3.16), that is

$$J^2 = J_1^2 + J_2^2 + J_3^2. \quad (3.40)$$

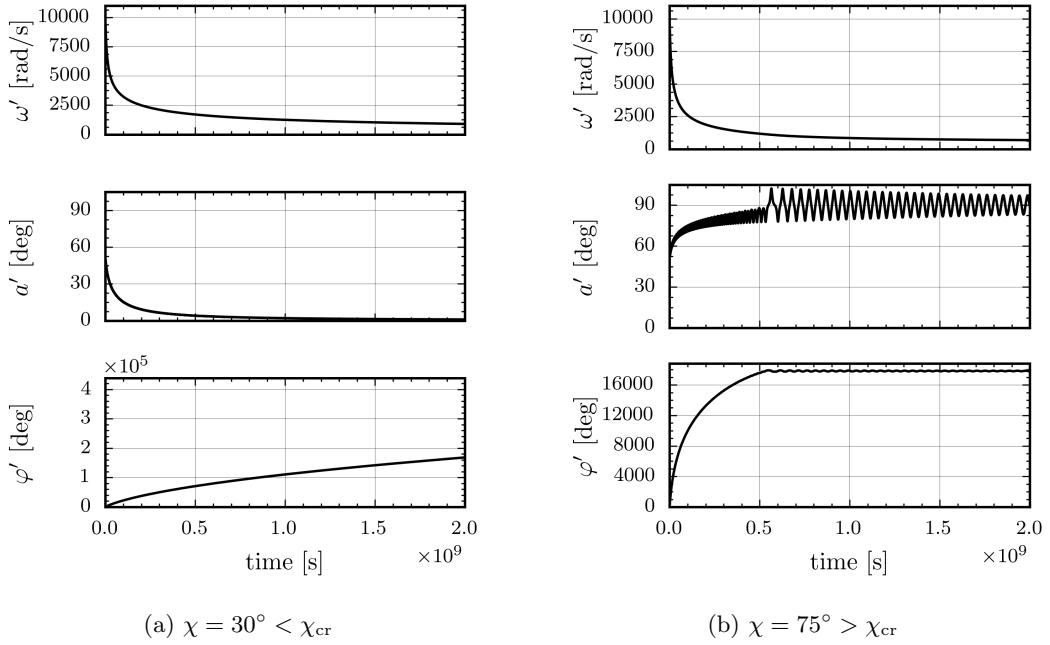


Figure 3.11: Plot of the spherical coordinates of the spin-vector for neutron star A. We note that for the $\chi = 75^\circ$ solution, eventually $a' \rightarrow 90^\circ$ as the oscillations damp.

In analogue to the geometric argument of Section 3.3.1, the first equation, the conservation of momentum, describes a *triaxial* ellipsoid with semi axes given by $\sqrt{2\lambda_- E}$, $\sqrt{2I_0 E}$, and $\sqrt{2\lambda_+ E}$, while the conservation of energy still describes a sphere of radius J . In order to satisfy both conservation equations, the sphere and ellipsoid must intersect bounding the radius of the sphere by

$$2E\lambda_- < J^2 < 2E\lambda_+. \quad (3.41)$$

In the case of a biaxial moment of inertia, we find that the solution for which both conservation laws are satisfied corresponded to circles around the \hat{z} axis. Now that we have a triaxial moment of inertia tensor (due to the addition of the anomalous torque), the solutions correspond to the intersection of the triaxial ellipsoid with the sphere, for which there are three distinct types of intersections as drawn in Figure 3.12.

Each of these types of intersection can be understood through the orderings of the radius of the sphere and the sizes of the three semi-axes of the ellipsoid:

- (a) $2E\lambda_- < J^2 < 2EI_0$. In this region, curves form two sets of complete loops around the e_1 axis. As $J^2 \rightarrow 2E\lambda_-$ the intersection loops will close up about the e_1 axis.
- (b) $J^2 = 2EI_0$. This is a special case in which the intersection forms two closed ellipses
- (c) $2EI_0 > J^2 > 2E\lambda_+$. In this region, curves form two sets of complete loops around the e_3 axis. As $J^2 \rightarrow 2E\lambda_+$ the intersection loops will close up about the e_3 axis.

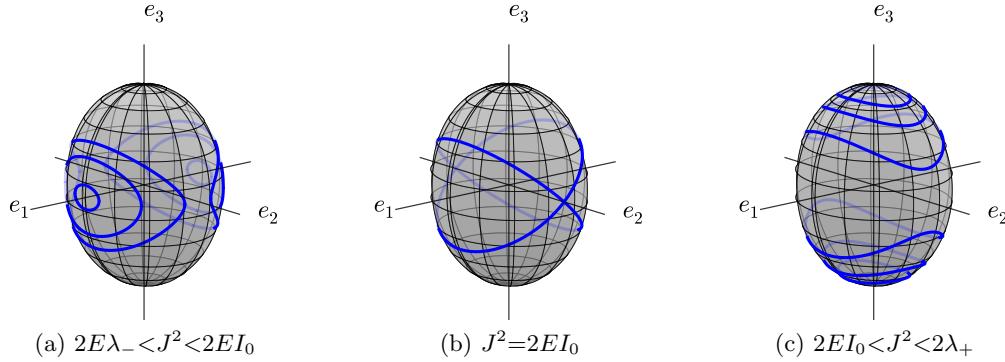


Figure 3.12: Intersections of the ellipsoid and sphere defined in Eqn. (3.39) and Eqn. (3.40) respectively for three different values of M at fixed E .

Now that we have understood the possible types of solution via the intersections of the sphere and ellipsoid, we can understand the cause of the unusual behaviour in Figure 3.11(b). First, we note that all simulations begin in the same initial state. As previously observed, changing χ through the critical value changes the sign of \dot{A} and hence whether the sphere will grow or shrink with respect to the ellipsoid. So for $\chi = 30^\circ$, the sphere grows with respect to the ellipsoid ($\dot{A} > 0$) ending up with the loops closing up about the e_3 axis. For $\chi = 75^\circ$ the sphere shrinks with respect to the ellipsoid ($\dot{A} < 0$) ending up with the loops closing up about the e_1 axis. The two end states are then the (a) and (c) pictures in Figure 3.12.

If they both begin from the same initial configuration, then one of them must pass through the special $J^2 = 2EI_0$ state corresponding to Figure 3.12(b). When this happens, we will observe a change in the axis about which the solution precesses.

Both sets of solution start with the solution precessing about the e_3 axis (φ monotonically increasing) corresponding to Figure 3.12(c). In Figure 3.11(b), the discontinuity is in fact exactly the point when the solution changes the axis of precession.

To better understand this, we plot the data from the $\chi = 75^\circ$ simulation (Figure 3.11(b)) in Figure 3.13. Firstly in (a), we project the spherical coordinates onto the unit sphere and choose three short sections of data. This shows that at early times (in blue) the solution precesses about the e_3 axis. During the discontinuity (in red) the solution shows similarities with the $J^2 = 2EI_0$ case from Figure 3.12. Finally, at late times, in black, the solution precesses about the e_1 axis. In Figure 3.13(b) we also plot the solution in 3D during the apparent discontinuity. This again shows that the solution goes through a change of the axis about which it precesses.

3.4.4 Neutron star B in region $\tau_S > \tau_P > \tau_A$

Solutions for neutron star B take on a new significance having included the anomalous torque – we can now make a comparison with the work of Melatos [122]. In this work,

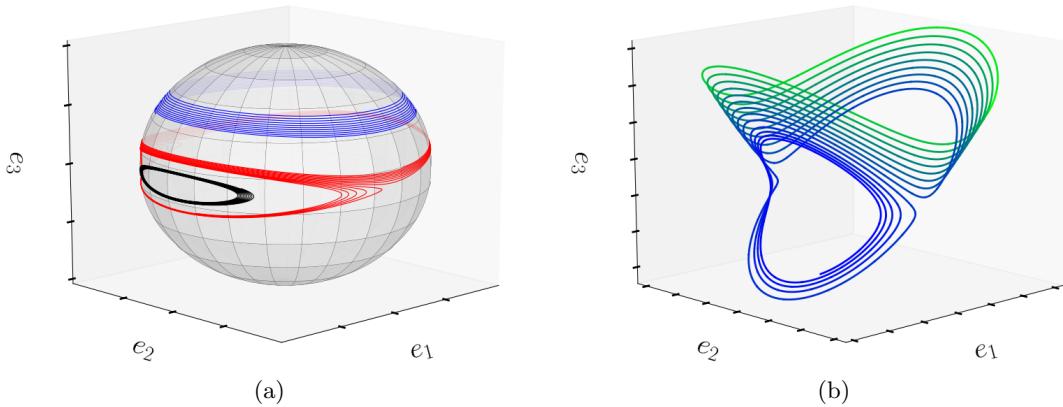


Figure 3.13: Numerical solutions for neutron star A with $\chi = 75^\circ$ plotted in 3D showing the point where the spin-vector changes the axis about which it precesses. Figure (a) shows the evolution of the spin-vector on the unit sphere at three distinct time steps: blue marks the time before the discontinuity, red is during the discontinuity, and black is some time after. Figure (b) shows the complicated behaviour of the spin-vector during the discontinuity, the changing colour is used to illustrate the evolution with time, progressing from green to blue.

Melatos found that when the precession timescale and anomalous torque timescales are comparable, the solutions involve *persistent precession*. Specifically, this is defined as the polar angle a tending to a constant, non-zero value with the nutation amplitude (oscillations in a) either decaying or remaining constant. Because the spin-vector has not aligned with a principal axis of the moment of inertia, this was interpreted as the spin-vector undergoing persistent precession.

We simulate neutron star B with the properties listed in Table 3.3 and a magnetic inclination angle $\chi = 75^\circ$. In Figure 3.14(a) and Figure 3.14(b) we plot the spherical components of the spin-vector in the body frame and effective body frame respectively. Figure 3.14(a), the components in the body frame, confirms the finding by Melatos that $a \neq 90^\circ$ at the end of the simulation.

However, transforming to the effective body frame in Figure 3.14(b), we find that the spin-vector aligns with the principal axis of the effective moment of inertia tensor ($a' \rightarrow \pi/2$). Therefore, the persistent precession angle found by Melatos is precisely the angle β . The spin-vector has aligned with the principal axis of the effective moment of inertia tensor associated with the smallest eigenvalue. Since both the polar angle a' and azimuthal angle φ' are constant (this is also true in the body frame), we argue that the angular velocity is constant and this is not persistent precession: the spin-vector always aligns with the effective body frame and in doing so the precession of the body is damped.

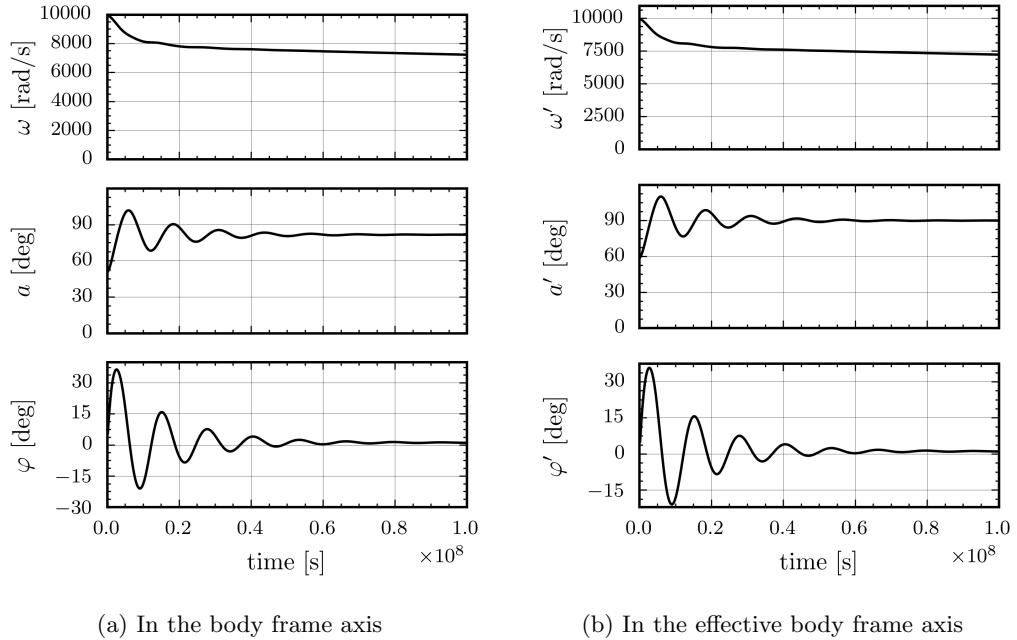


Figure 3.14: Plot of the spherical coordinates of the spin-vector for neutron star B. Crucially, at late times the polar angles a and a' differ by a factor 8.4° , the value of β in Table 3.3.

3.4.5 Neutron star C in region $\tau_P > \tau_S > \tau_A$

In Figure 3.15, for all values of χ , the spin-vector is found to align with the effective body frame e_1 axis. In this limit where $\epsilon_A \gg \epsilon_I$, from Figure 3.9, we see that the e_1 axis aligns with the magnetic dipole. This agrees with the results for a spherical star in which $\epsilon_I \rightarrow 0$. The addition of the anomalous torque introduces oscillations on the anomalous torque timescale.

3.5 Conclusion

In this chapter, we have studied a numerical model comprising a rigid body supporting a biaxial strain and acted on by the Deutsch [54] dipole torque. In agreement with Melatos [122], we confirm the analytic solutions of Goldreich [69]: when $\tau_P \ll \tau_S$ the inclination of the spin-vector a either damps or grows on the spin-down timescale depending on whether χ is smaller or larger than $\chi_{\text{cr}} \approx 55^\circ$. Moreover, we provide a new interpretation of this critical value as a tipping point between the rates of change of energy and angular momentum.

When $\tau_P \ll \tau_S$ we find that the spin-vector aligns with the magnetic dipole; this also agrees with the numerical and analytic solutions found by Melatos [122]. When including the anomalous torque, in the regime where $\tau_P \sim \tau_S$, we confirm the existence of so-called persistent precession solutions found by Melatos [122] in which the spin-vector does not align with either the principal axis of the moment of inertia tensor or the magnetic

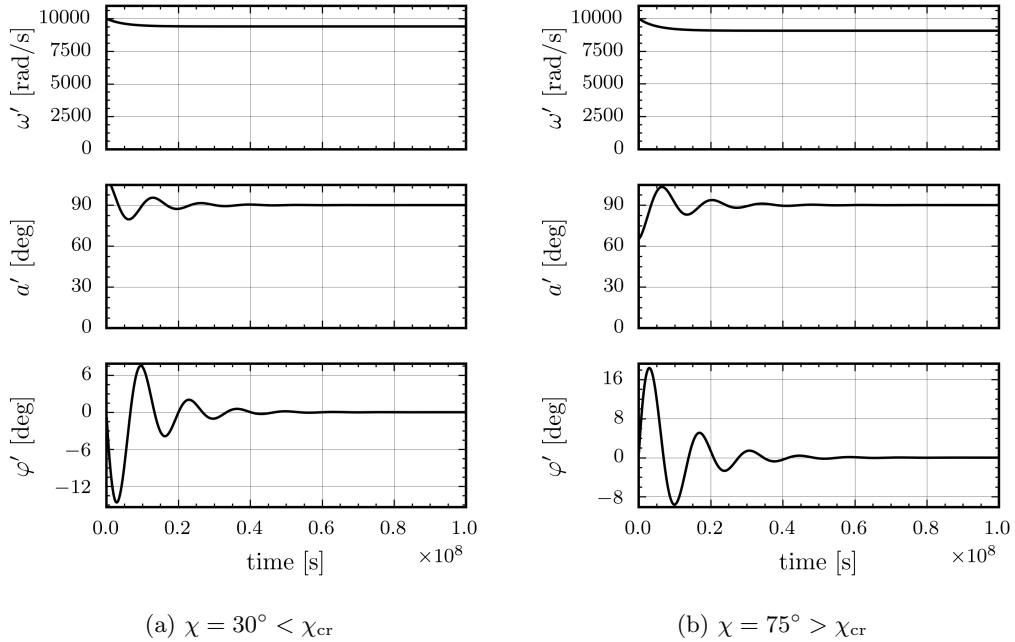


Figure 3.15: Plot of the spherical coordinates of the spin-vector for neutron star C.

dipole. However, we provide a new insight into this by showing that for all orderings of the three timescales, the spin-vector aligns with the principal axes of the effective moment of inertia tensor (defined in Section 3.4.1). In this sense, there is no persistent precession.

By studying solutions with and without the anomalous torque, we conclude that the anomalous torque is not important for realistic systems since it amounts to nothing more than a small shift in the axes about which that spin-vector precesses. For this reason, the solutions found by Goldreich [69] is the most relevant for real pulsars and it will be this that we use in the rest of this thesis.

Appendix 3.A Considerations for the timescales

We have used the ordering of the three timescales at $\omega_0 = \omega(t = 0)$ to categorise the results. However, clearly the magnitude of the spin-vector will decrease and the different dependencies on the spin-vector could cause a ‘crossing’ of the timescales producing unexpected results. Writing the timescales as

$$\tau_P(t) = \frac{2\pi}{\epsilon_I \omega(t)}, \quad (3.42)$$

$$\tau_A(t) = \frac{2\pi}{\epsilon_A \omega(t)}, \quad (3.43)$$

$$\tau_S(t) = \frac{3c}{2R} \frac{2\pi}{\epsilon_A \omega(t)^2}, \quad (3.44)$$

only τ_P could ‘cross’ with the other two timescales since the spin-down and anomalous timescales obey $\tau_S > \tau_A$ provided that

$$\omega(t) > \frac{3c}{2R}. \quad (3.45)$$

This condition is satisfied by setting the initial rotational spin frequency at less than $\omega_0 \sim 10^4$ [rad/s], that is the star should not break special relativity. At later times, the spin frequency will decay and so this condition is still satisfied.

For the anomalous torque and precession timescale crossings, two cases exists: either the deformations are such that the spin-down timescale is initially larger than the precession timescale (region A and B) or, as in region C, the precession timescale is larger than the spin-down timescale. Let us consider the first case:

$$\tau_S > \tau_P \Rightarrow \omega(t) < \frac{3c}{2R} \frac{\epsilon_I}{\epsilon_A}. \quad (3.46)$$

In this particular state, $\epsilon_I > \epsilon_A$ and so, as in the previous case, while the spin-vector decays this inequality is always satisfied and the orderings remain the same. In the other case we have

$$\tau_P > \tau_S \Rightarrow \omega(t) > \frac{3c}{2R} \frac{\epsilon_I}{\epsilon_A}. \quad (3.47)$$

In this case, it is possible for the timescales to cross at $\omega_{\text{cr}} = \frac{3c}{2R} \frac{\epsilon_I}{\epsilon_A}$. For pulsar C, this corresponds to a rotational spin frequency of 0.9 [rad/s], which is four orders of magnitude smaller than the initial frequency. For this reason, we can rule out this crossing of the timescales as an important factor in calculations.

Chapter 4

Modelling observations of precessing pulsars

In Section 3, we investigated the role of precession under an EM torque in the rotating frame of the star. From this we learnt that the Goldreich [69] analytic solutions, which ignore the anomalous component of the Deutsch [54] torque, are the relevant solutions for the precession of a torqued pulsar. We will now model the effect of precession on the ‘observable features’ of a pulsar, by which we mean the phase residual, spin-down rate, and pulse profile. This will require us to view the pulsar in the inertial frame of an observer and model the data collection methods of pulsar astronomy as part of a predictive model of pulsar precession.

Modelling observations of precessing pulsars has a history dating back to Ruderman [149] and Chiuderi and Occhionero [41] when they were attempting to explain a small amplitude ‘wobble’ in the frequency of the Crab pulsar [147, 146]. Since then, new observations have prompted many workers in the field to develop predictive models and in the course of this chapter we will discuss their contributions when relevant. The novel material presented here is in numerically solving the Euler rigid-body equations under the full Deutsch [54] torque in addition to solving for the Euler angles which allows us to make direct predictions of the observable features measured by observers. Moreover, we will also model the data collection mechanisms such as the *observer-method* to measure the evolution of the spin-down rate. We will compare this numerical model against the analytic results from the literature, in particular we will use the work of Jones and Andersson [90] as it is the most complete work on the subject. By comparing with a numerical model, we can gain confidence in the use of analytic solutions. For example, in Chapter 5 we will make use of the analytic solution for the spin-down rate of a precessing pulsar which we will derive in Section 4.5.1; this derivation neglects the anomalous torque discussed in Section 3, but we can be confident in its predictions because in Section 4.5.2 we compare the analytic approximation against numerical solutions calculated with the full torque.

This chapter is organised in the following way. In Section 4.1 we will introduce the idea of the Euler rotation angles and how to formulate a predictive model for precessing pulsars

in the inertial frame. In Section 4.2 we compare some typical numerical solutions to analytic solutions in the torque free case and then illustrate the effect of the full Deutsch [54] torque on the evolution of the Euler angles. Before discussing the observable features of pulsars, we need some intuition to understand the results, so in Section 4.3 we give a detailed account of the motion of the magnetic dipole of a freely precessing pulsar. The observable features of a precessing pulsar are discussed in Section 4.4, Section 4.5, and Section 4.6 for the timing residuals, spin-down rate, and shape of pulsations respectively. Finally, in Section 4.7 we give some preliminary findings on a hybrid model of precession and magnetospheric switching before concluding in Section 4.8.

4.1 Rotating into the inertial frame

4.1.1 Euler rotation matrices

The Euler rigid-body equation, introduced in Chapter 3, is defined in the rotating frame of the star. To discuss results in the inertial frame of an observer, we need to transform the solutions of Euler's rigid-body equations into the inertial frame. An efficient way to do this is to determine the three Euler angles which transform the rotating frame axes, denoted by (x', y', z') , to the inertial frame axis, denoted by (x, y, z) . In particular, we define three rotation matrices

$$B = \begin{bmatrix} \cos \psi & \sin \psi & 0 \\ -\sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (4.1)$$

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & \sin \theta \\ 0 & -\sin \theta & \cos \theta \end{bmatrix}, \quad (4.2)$$

$$D = \begin{bmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (4.3)$$

Then their product defines the Euler angle rotation matrix

$$R = BCD. \quad (4.4)$$

In index notation the components are

$$R_a^b = \begin{bmatrix} \cos \psi \cos \phi - \cos \theta \sin \phi \sin \psi & \cos \psi \sin \phi + \cos \theta \cos \phi \sin \psi & \sin \psi \sin \theta \\ -\sin \psi \cos \phi - \cos \theta \sin \phi \cos \psi & -\sin \psi \sin \phi + \cos \theta \cos \phi \cos \psi & \cos \psi \sin \theta \\ \sin \theta \sin \phi & -\sin \theta \cos \phi & \cos \theta \end{bmatrix}. \quad (4.5)$$

Here we are using the (ϕ, θ, ψ) Euler angle parameterisation as described in §35 of Landau and Lifshitz [100]; a diagram of how these angles are defined is given in Figure 4.1. This matrix, given values for the Euler angles, transforms from the rotating frame to the fixed inertial system of coordinates. So for A_b^{rot} , a vector defined in the inertial frame, in the

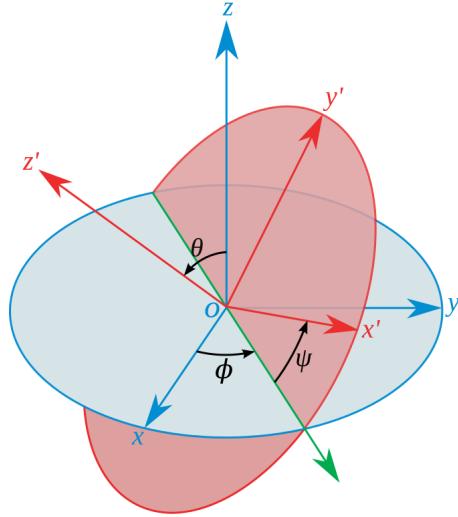


Figure 4.1: The Euler rotation angles between the rotating frame (x', y', z') and the inertial frame (x, y, z) . Image courtesy of Brits [36].

rotating frame the vector has components given by

$$A_a^{\text{in}} = R^b{}_a A_b^{\text{rot}}. \quad (4.6)$$

4.1.2 Evolution of the Euler angles

The Euler angles themselves will evolve with time. To calculate this, we express the components of the spin-vector Ω_a along the moving axes (x', y', z') . As shown by Landau and Lifshitz, this results in a set of three ODEs which are coupled both to the other Euler angles and the components of the spin-vector. To have solutions both for the motion of the spin-vector in the rotating frame and the Euler angles we need to solve all six ODEs together. To be specific, the three Euler rigid-body equations for a biaxial body with a moment of inertia as defined in Eqn. (2.8) are given by

$$\dot{\Omega}_x = \frac{T_x}{I_0} - \epsilon_I \Omega_y \Omega_z, \quad (4.7)$$

$$\dot{\Omega}_y = \frac{T_y}{I_0} + \epsilon_I \Omega_x \Omega_z, \quad (4.8)$$

$$\dot{\Omega}_z = \frac{T_z}{I_0(1 + \epsilon_I)}. \quad (4.9)$$

The rearranged set of Euler angle ODEs from Equation (35.1) of Landau and Lifshitz [100] are given by

$$\dot{\phi} = \frac{\Omega_x \sin \psi + \Omega_y \cos \psi}{\sin \theta}, \quad (4.10)$$

$$\dot{\theta} = \Omega_x \cos \psi - \Omega_y \sin \psi, \quad (4.11)$$

$$\dot{\psi} = \Omega_z - \dot{\phi} \cos \theta, \quad (4.12)$$

where $\epsilon_I = \Delta I/I_0$ is the deformation of star. This set of six coupled ODEs can be solved numerically using a time stepper; we will use the `rkf45` stepper provided by GSL [72]. Solutions give the components of the spin-vector in the rotating frame and the evolution of the Euler angles which can be used to transform rotating frame quantities into the inertial frame. Unlike the results of Section 3, numerical solutions to these ODEs require the fast spin frequency to be resolved and hence require greater computing time.

4.1.3 Initial conditions

Solving the rigid-body equations, as in Section 3, we set $\Omega_a(t=0)$ to lie in the $x' - z'$ plane at an angle a_0 to the z' axis. This gives a set of three initial conditions

$$\Omega_x = \omega_0 \sin(a_0), \quad \Omega_y = 0, \quad \Omega_z = \omega_0 \cos(a_0), \quad (4.13)$$

where ω_0 is the initial magnitude of the spin-vector, recalling that we define $\omega = |\Omega_a|$.

For the Euler angle equations, Eqn. (4.11) to Eqn. (4.10), the initial conditions need to be chosen carefully so that the result can be meaningfully interpreted. In particular, we need to be sure we understand how the inertial frame is orientated. Let us note that the angular momentum in the two frames are related by

$$J_a^{\text{in}} = R_b^b J_b^{\text{rot}}. \quad (4.14)$$

where R_b^b is defined in Eqn. (4.5).

We have already set the initial condition on Ω_a in Eqn. (4.13). Therefore the initial angular momentum in the rotating frame is given by

$$J_a^{\text{rot}}(t=0) = I_a^b \Omega_b(t=0). \quad (4.15)$$

If we set an initial condition on the angular momentum in the inertial frame J_a^{in} , then Eqn. (4.5) uniquely defines the initial Euler angles. We choose to set the initial angular momentum in the inertial frame to lie along the inertial z axis such that

$$J_a^{\text{in}}(t=0) = |J_a^{\text{in}}| \hat{z}. \quad (4.16)$$

The magnitude of the angular momentum is

$$|J_a| = |I_a^b \Omega_b| = \omega_0 I_0 \sqrt{\sin^2 a_0 + \cos^2 a_0 (1 + \epsilon_I)^2}. \quad (4.17)$$

Using the inverse of Eqn. (4.14) and substituting in Eqn. (4.15) and Eqn. (4.16) we have

$$\begin{bmatrix} \sin a_0 \\ 0 \\ (1 + \epsilon_I) \cos a_0 \end{bmatrix} = \sqrt{\sin^2 a_0 + \cos^2 a_0 (1 + \epsilon_I)^2} \begin{bmatrix} \sin \psi_0 \sin \theta_0 \\ \cos \psi_0 \sin \theta_0 \\ \cos \theta_0 \end{bmatrix}. \quad (4.18)$$

This gives us three equations for two unknowns. Our choice to set J_a^{in} along the z axis leaves the initial value of ϕ as a free variable. We set $\phi(t=0) = 0$ without loss of

generality. Rearranging the third component of Eqn. (4.18) yields

$$\theta_0 = \arccos \left[\frac{(1 + \epsilon_I) \cos a_0}{\sqrt{\sin^2 a_0 + (1 + \epsilon_I)^2 \cos^2 a_0}} \right]. \quad (4.19)$$

In the limit $\epsilon_I \ll 1$, it can be seen that $\theta_0 \approx a_0$. For ψ_0 , we rearrange the first component of Eqn. (4.18) to give

$$\sin \psi_0 = \frac{1}{\sin \theta_0} \frac{\sin a_0}{\sqrt{\sin^2 a_0 + (1 + \epsilon_I)^2 \cos^2 a_0}}. \quad (4.20)$$

To simplify the first factor, we use the identity $\sin(\arccos(x)) = \sqrt{1 - x^2}$ along with Eqn. (4.19) giving

$$\sin \theta_0 = \left[\frac{\sin^2 a_0}{\sin^2 a_0 + (1 + \epsilon_I)^2 \cos^2 a_0} \right]^{1/2}. \quad (4.21)$$

Inserting this into Eqn. (4.20) and rearranging we find that

$$\sin \psi_0 = \frac{\sin a_0}{(\sin^2 a_0)^{1/2}} \quad (4.22)$$

$$= \frac{\sin a_0}{|\sin a_0|} \quad (4.23)$$

$$= \text{sign}(a_0), \quad (4.24)$$

where by $\text{sign}(x)$ we mean the sign of x . Finally, the initial condition is given by

$$\psi_0 = \text{sign}(a_0) \frac{\pi}{2}. \quad (4.25)$$

We can check the sanity of this result by inserting it into the second component of Eqn. (4.18) and finding that it balances the left hand side.

In this section, we have defined the appropriate initial conditions for the system: in addition to Eqn. (4.13) defining the initial spin-vector, we set the angular momentum in the inertial frame to lie along the z axis and fix $\phi(t = 0) = 0$ without loss of generality. While the initial conditions on the spin-vector are arbitrary, if the initial Euler angle are not carefully defined then we do not have a meaningful interpretation for how the inertial frame is orientated.

4.2 Evolving the Euler angles

Having defined the coupled ODEs in Eqn. (4.7) to Eqn. (4.12) and appropriate initial conditions, we can now calculate the evolution of the spin-vector and Euler angles. The ultimate aim of this section is to go on to simulate the observable features. However, let us begin by considering the evolution of the simulation parameters (the components of Ω_a and the three Euler angles). We will consider first a torque free simulation in Section 4.2.1 and then add in the full torque in Section 4.2.2.

4.2.1 Torque free biaxial body

The set of ODEs given in Eqn. (4.7) to Eqn. (4.12) for a biaxial body free of torques have simple analytic solutions [100] given by

$$\theta(t) = \theta_0 \approx a_0, \quad (4.26)$$

$$\phi(t) = \dot{\phi}t + \phi_0 = \dot{\phi}t, \quad (4.27)$$

$$\psi(t) = \dot{\psi}t + \psi_0 = -\frac{\Delta I}{I_0} \dot{\phi}t + \text{sign}(a_0) \frac{\pi}{2}, \quad (4.28)$$

where in the second equalities we have inserted the initial conditions. Solution to the rigid-body equations were derived in Eqn. (2.11) to Eqn. (2.13), transforming these into the spherical coordinates of the spin-vector we have

$$|\Omega_a| = \omega_0, \quad (4.29)$$

$$a = a_0, \quad (4.30)$$

$$\varphi = \epsilon_I \omega_0 \cos(a_0)t, \quad (4.31)$$

where φ is the azimuthal angle of the spin-vector with respect to the $x' - z'$ plane. This set of analytic solutions gives us a method to verify our numerical solver with the appropriate initial conditions.

In Figure 4.2 we present the numerical solutions for some arbitrary values of the simulation parameters along with the analytic prediction of Eqn. (4.26) to Eqn. (4.31). This demonstrates ‘almost’ perfect agreement between the two: the components of Ω_a behave as expected; ϕ monotonically increases at the spin frequency; ψ decreases at the slower precession frequency; the polar angle θ should remain constant during this simulation, however, we find it varies fractionally by $\sim 10^{-11}$. This error is caused by the finite numerical precision when performing the subtraction in Eqn. (4.11).

4.2.2 Torqued biaxial body

To see the difference introducing the EM torque of Eqn. (3.5) makes to the solutions, in Figure 4.3 we repeat the simulation plotted in Figure 4.2 with both the anomalous and spin-down components of the EM torque. The most striking contrast is the ‘wobble’ in both θ and a ; on closer inspection one also finds this wobble in the other angles, while the magnitude of Ω_a both wobbles and undergoes a secular spin-down. To help illustrate this we have additionally plotted the derivatives $\dot{\phi}$ and $\dot{\psi}$ which were constant in the torque free case. These are calculated numerically, the numerical errors produce the observed noise.

These solutions show how, using numerical simulations, we can easily simulate the effect of the full Deutsch [54] torque. In the rest of this chapter, we will discuss how to model the features of a star observed by pulsar astronomers. This will allow us to change the physics of the star, i.e. by modifying the torque, and directly simulate the effect this has on the observable properties.

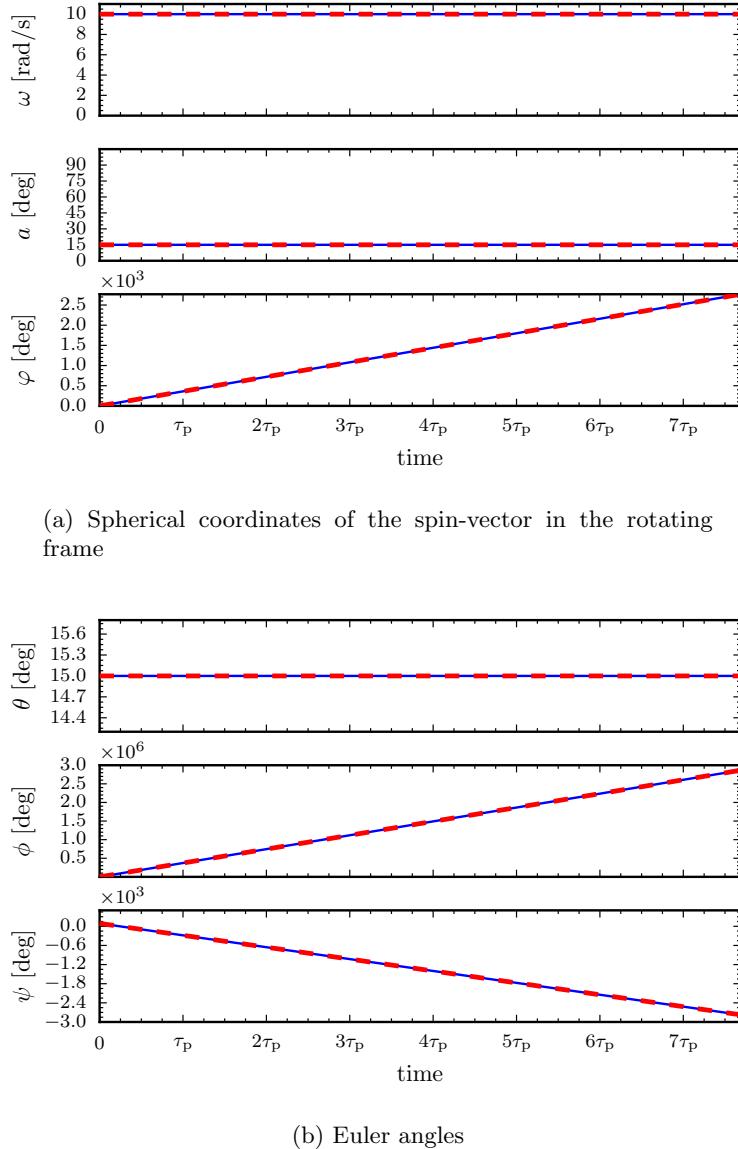
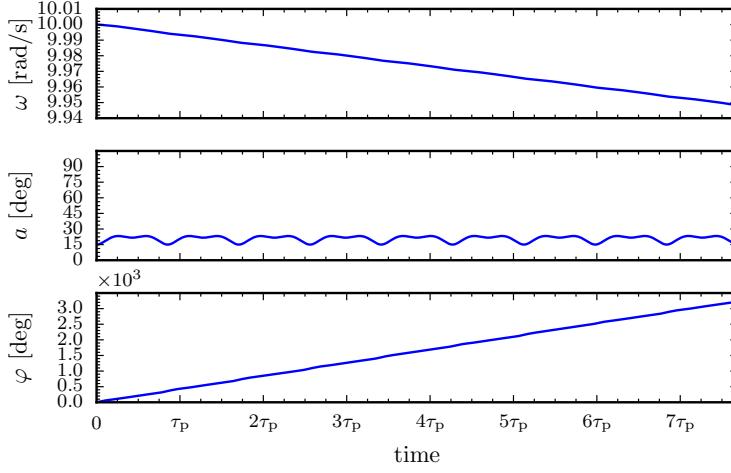
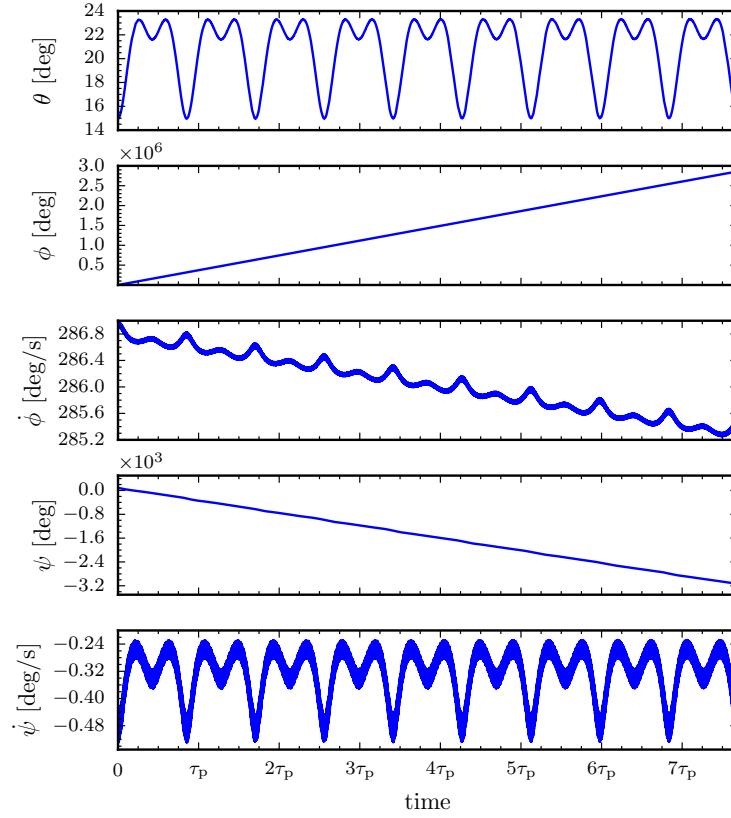


Figure 4.2: Solutions to the ODEs defined in Eqn. (4.7) to Eqn. (4.12) for a torque free biaxial star with a deformation of $\epsilon_I = 10^{-3}$. The red dashed line is the analytic calculation found by Jones and Andersson [90] as given in Eqn. (4.26) to Eqn. (4.31).



(a) Spherical coordinates of the spin-vector in the rotating frame



(b) Euler angles and selected derivatives

Figure 4.3: Solutions to the ODEs defined in Eqn. (4.7) to Eqn. (4.12) including the torque defined in Eqn. (3.5) for a biaxial body with $\epsilon_A = \epsilon_I/2$ and $\epsilon_I = 1 \times 10^{-3}$; the ratio of the precession timescale to the spin-down age is $\approx 7 \times 10^{-4}$. Note the x-axis is in units of τ_p , the initial precession period.

4.3 Precessing pulsars

Before we move to modelling the observable features such as phase-residuals, we need to build our intuition for precession. To do this, we will introduce the notion of a *reference plane* from which we can decompose the motion of vectors into rotations about cones.

4.3.1 The reference plane

Considering a biaxial body, let us define \hat{n}^a as the *deformation vector* (a unit vector pointing along the body's symmetry axis). Then the moment of inertia tensor can be written compactly as

$$I^a_b = I_0 \delta^a_b + \Delta I \hat{n}^a \hat{n}_b, \quad (4.32)$$

where δ^a_b is the Kronecker delta. The principal moments are given by $I_0 = I_{xx} = I_{yy}$ and $I_{zz} = I_0 + \Delta I$ such that the body is biaxial with \hat{n}_a lying along the \hat{z} axis. As such, the angular momentum is given by

$$J_a = I_0 \Omega_a + \Delta I \Omega_z \hat{n}_a. \quad (4.33)$$

As found by Pines and Shaham [136], this shows that the three vectors J_a , Ω_a , and \hat{n}_a are coplanar in the so-called reference plane as shown in Figure 4.4. In this figure, we

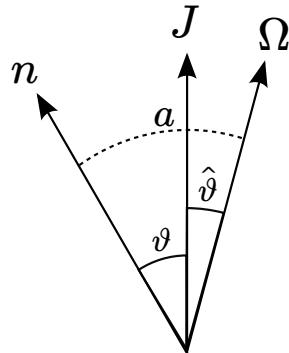


Figure 4.4: The reference plane containing the spin-vector Ω , the angular momentum vector \mathbf{J} and the deformation vector \mathbf{n} .

have defined the angle between angular momentum and \hat{n}_a as θ , conventionally referred to as the *wobble-angle* of precession. We have also given a , the angle between the \hat{z} -axis and the spin-vector as defined in Section 3.

Following Jones and Andersson [90], let us now decompose the angular velocity into the Euler angle components along the angular momentum and symmetry axis (these are the same Euler angles as defined in Section 4.1.1)

$$\Omega_a = \dot{\phi} \hat{J}_a + \dot{\psi} \hat{n}_a, \quad (4.34)$$

where \hat{J}_a is a unit vector pointing along the angular momentum. Then, substituting this into Eqn. (4.33) we find that

$$J_a = I_0 \dot{\phi} \hat{J}_a + \left(I_0 \dot{\psi} + \Delta I \Omega_z \right) \hat{n}_a. \quad (4.35)$$

Comparing components, we have that

$$|J_a| = I_0 \dot{\phi}, \quad (4.36)$$

and

$$\dot{\psi} = -\frac{\Delta I}{I_0} \Omega_z. \quad (4.37)$$

To relate the rates of change of the two angles, we can use Eqn. (4.12), giving

$$\dot{\psi} = -\frac{\Delta I}{I_0 + \Delta I} \dot{\phi} \cos \theta \approx -\frac{\Delta I}{I_0} \dot{\phi} = -\epsilon_I \dot{\phi} \quad (4.38)$$

where we assumed $\Delta I \ll I_0$ and $\theta \ll 1$ in the last step.

We can relate both of these frequencies to the angular frequency $\omega = |\Omega_a|$. First, from the cosine rule we have that

$$|\Omega_a|^2 = \dot{\phi}^2 + \dot{\psi}^2 - 2\dot{\phi}\dot{\psi} \cos \theta. \quad (4.39)$$

and so rearranging and neglecting terms of order ϵ_I^2 , we have that

$$\dot{\phi} \approx \frac{\omega}{\sqrt{1 - 2\epsilon_I \cos \theta}}. \quad (4.40)$$

Since $\epsilon_I \ll 1$, we know that $\dot{\phi}$ is approximately the fast spin frequency of the star ω . Then from Eqn. (4.38), we see that $\dot{\psi} \ll \omega$ is a much slower frequency, the precession frequency.

We will now discuss the effect of free precession on the motion of \hat{m}_a , the magnetic dipole. We will do this initially by studying the evolution of the dipole in the inertial frame via the Euler angle rotations and then we will return to this reference frame picture in Section 4.3.3 to provide a deeper intuitive understanding.

4.3.2 Dynamics of the magnetic dipole

Numerical solutions to Eqn. (4.7) to Eqn. (4.12) allow us to calculate the motion of any quantity in the inertial frame from which the neutron star is observed. This can be used, for example, to calculate the motion of the spin-vector as seen by an observer for any arbitrary torque. However, pulsar astronomers observe the pulsar through the pulsations of EM emission. If this emission is collinear with the dipole, then it points along the unit vector of the magnetic dipole \hat{m}_a . Therefore, we are particularly interested in the motion of \hat{m}_a in the inertial frame. This system was first studied in the follow

way by Bisnovatyi-Kogan et al. [30], but here we follow the treatment by Jones and Andersson [90].

In the rotating frame, we set \hat{m}_a to lie at an angle χ to the z' axis with unit vector $[\sin(\chi), 0, \cos(\chi)]$ without loss of generality. Using Eqn. (4.5), the Euler rotation matrix, we can transform to the inertial frame; the components of the magnetic dipole in the inertial frame are

$$\hat{m}_a = \begin{bmatrix} \cos \phi \cos \psi \sin \chi - \sin \phi \cos \theta \sin \psi \sin \chi + \sin \phi \sin \theta \cos \chi \\ \sin \phi \cos \psi \sin \chi + \cos \phi \cos \theta \sin \psi \sin \chi - \cos \phi \sin \theta \cos \chi \\ \sin \theta \sin \psi \sin \chi + \cos \theta \cos \chi \end{bmatrix}. \quad (4.41)$$

We define two angles Φ and Θ which describe the polar and azimuthal angles of \hat{m}_a in the inertial frame. From Eqn. (4.41) the azimuthal angle is given by

$$\Phi = \arctan\left(\frac{\hat{m}_y}{\hat{m}_x}\right) = \phi - \frac{\pi}{2} + \arctan\left[\frac{1}{\cos \theta}\left(\frac{\cos \psi \tan \chi}{\tan \theta - \sin \psi \tan \chi}\right)\right], \quad (4.42)$$

while the polar angle is

$$\Theta = \arccos(\hat{m}_z) = \arccos(\sin \theta \sin \psi \sin \chi + \cos \theta \cos \chi). \quad (4.43)$$

Eqn. (4.42) and Eqn. (4.43) are exact results and we will use them later in this chapter to model the observable features such as the phase residuals and spin-down rate. Specifically, given a numerical solution to our system of coupled ODEs, i.e. the evolution of the spin-vector and three Euler angles, we can use these two equations to describe the evolution of the magnetic dipole orientation in the inertial frame. Before we do this, let us now build some feeling for the dynamics of the magnetic dipole in the inertial frame in the torque free case.

Using our numerical solution, we simulate a star without any EM torque. The resulting solutions for the Euler angles can then be substituted into Eqn. (4.43) to give the evolution of Θ . This is done for three choices of θ and χ in Figure 4.5. From this figure, we see that the polar angle is modulated at the precession period, τ_P . In the next section, we will provide a deeper understanding of the size of modulations and the significance in the choice of θ and χ .

The observed spin frequency of a pulsar is the rate with which observers measure the observed pulsations. Usually we think of this as exactly equivalent to the rotation frequency $\omega/2\pi$. However, for a precessing star the observed frequency is not given by $\omega/2\pi$ due to the additional motions of precession. Instead, the observer would fit a timing model to the TOAs as the pulse passes through the plane containing the observer and the angular momentum vector, the phase of which is Φ . So the frequency measured would be given by the first derivative of Φ :

$$\dot{\Phi} = \dot{\phi} + \frac{\sin \chi \left(\dot{\psi}(\cos \theta \sin \chi - \sin \psi \sin \theta \cos \chi) + \dot{\theta} \cos \psi (\cos \theta \cos \chi - \sin \psi \sin \theta \sin \chi) \right)}{(\sin \theta \cos \chi - \cos \theta \sin \psi \sin \chi)^2 + \cos^2 \psi \sin^2 \chi}. \quad (4.44)$$

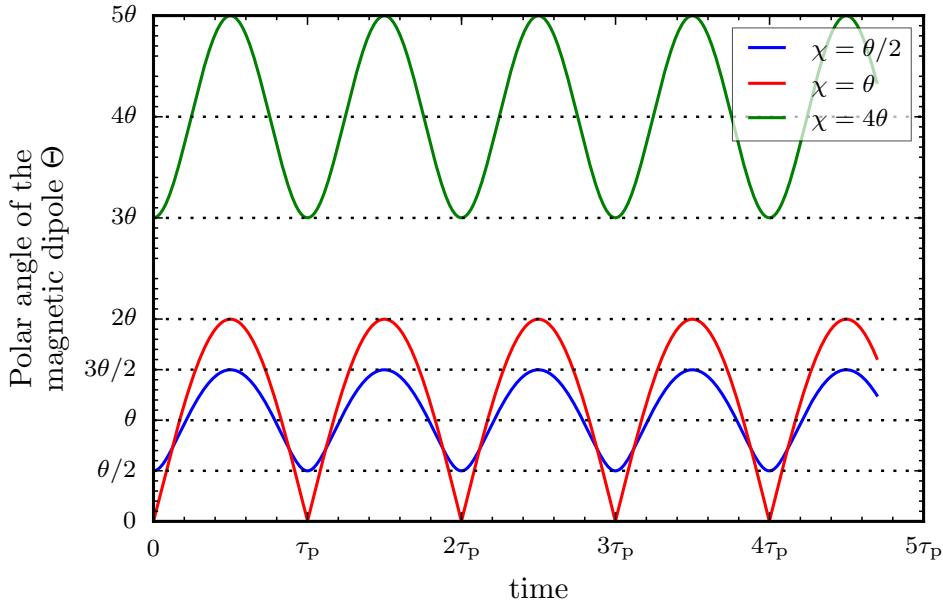


Figure 4.5: The effect of free precession on the polar angle of the dipole \hat{m}_a as given by Eqn. (4.43) for three choices of χ and θ .

This is the *instantaneous electromagnetic frequency*; an observer will measure the time averaged value of $\dot{\Phi}$ as the ‘spin frequency’ of the star.

In Figure 4.6 we plot the frequency modulations for the three choices of θ and χ used in Figure 4.5. Again, we see modulation at the precession period; in the following section, we will provide a deeper understanding of these modulations and our choices of θ and χ .

4.3.3 Understanding the dynamics of the magnetic dipole

We will now return to the reference plane picture discussed in Section 4.3.1. In particular, we provide a detailed description of the motion of \hat{m}_a as seen by the observer and use this to explain the modulations seen in Figure 4.5 and Figure 4.6.

Firstly, let us state our assumption that EM radiation is emitted along the dipole axis \hat{m}_a such that it is this axis from which an observer sees pulsations. From the decomposition of the spin-vector in Eqn. (4.34), the motion of the magnetic dipole, a vector fixed in the rotating frame, when viewed in the inertial frame, can be understood as the superposition of two rotations. If we hold ϕ fixed, then \hat{m}_a rotates in a cone of half-angle χ about the symmetry axis \hat{n}_a at the slow precession angular frequency $\psi \approx -\epsilon_I \omega$ and we call this the *precession cone*. On the other hand, if we hold ψ fixed, then \hat{n}_a rotates in a cone of half-angle θ about J_a at the angular frequency $\dot{\phi} \approx \omega$. Still holding ψ fixed, since \hat{m}_a is necessarily always at a fixed angle χ to \hat{n}_a , it too will be swung in a cone of half-angle Θ (as given in Eqn. (4.43)) about the angular momentum vector with angular frequency $\dot{\phi}$; let us define this to be the *dipole cone*. Now, ψ evolves at a slower

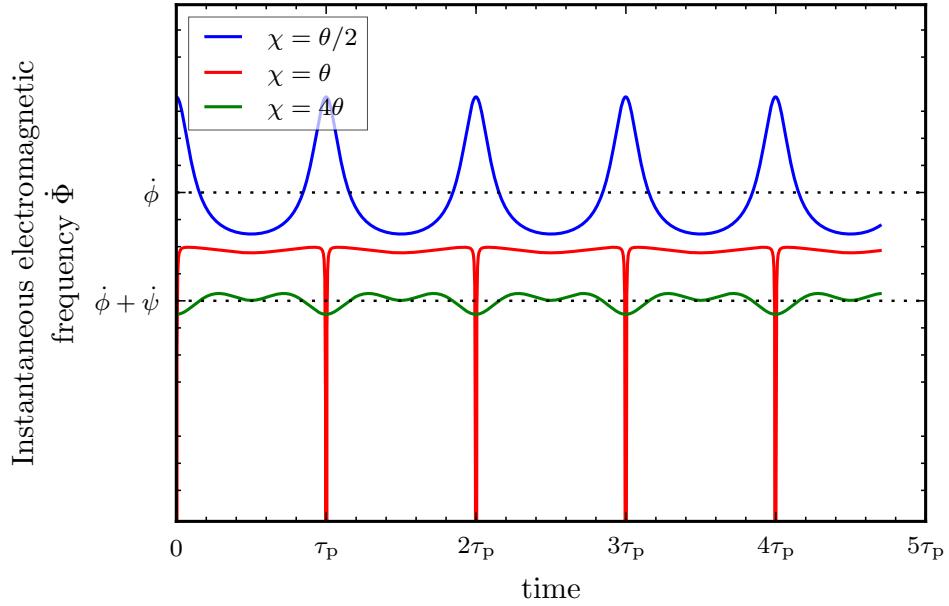


Figure 4.6: The effect of free precession on the instantaneous electromagnetic frequency given by Eqn. (4.44) for three choices of χ and θ .

rate to ϕ by a factor ϵ_I , moreover, the negative relation between them in Eqn. (4.38), indicates that the precession cone counter-rotates with respect to the dipole cone. If we consider the effect of both rotations together and their effect on the dipole cone, we see that the half-angle of the dipole cone Θ is modulated by precession (see Figure 4.5): as \hat{m}_a rotates about the dipole cone, the cone gradually opens and closes at the precession period. The rate at which \hat{m}_a rotates about the dipole cone is given by $\dot{\Phi}$, as given by Eqn. (4.44). Similarly this is also modulated by precession as shown in Figure 4.6.

The resulting motion of the dipole cone is best understood by considering three choices of χ and θ : $\chi < \theta$, the special case $\chi = \theta$, and $\chi > \theta$. We illustrate the projection of the precession cone swept out by \hat{m}_a about \hat{n}_a (in red) and the cone swept out by \hat{n}_a about J_a onto the reference plane for the three orderings of χ and θ in Figure 4.7. Note that the dipole cone is not shown here.

Let us now describe the evolution of the dipole cone for the three cases in Figure 4.7. We will discuss in particular the choices of χ and θ which were used in Figure 4.5 and Figure 4.6 for the variations in polar angle and instantaneous electromagnetic frequency respectively.

- The $\chi < \theta$ case ($\chi = \theta/2$): the precession cone is narrow and does not extend over the angular momentum vector. The polar angle Θ of the dipole cone oscillates periodically between $\theta + \chi$ and $\theta - \chi$ during a precession cycle, as shown in Figure 4.5. The spin frequency $\dot{\Phi}$, as shown in Figure 4.6, has an average value of $\dot{\phi}$ and oscillates about this value. Comparing with the Θ variations demonstrates these oscillations are locked in phase with the rotation of \hat{m}_a in the precession cone. Recalling that the precession cone counter-rotates with respect to the spin

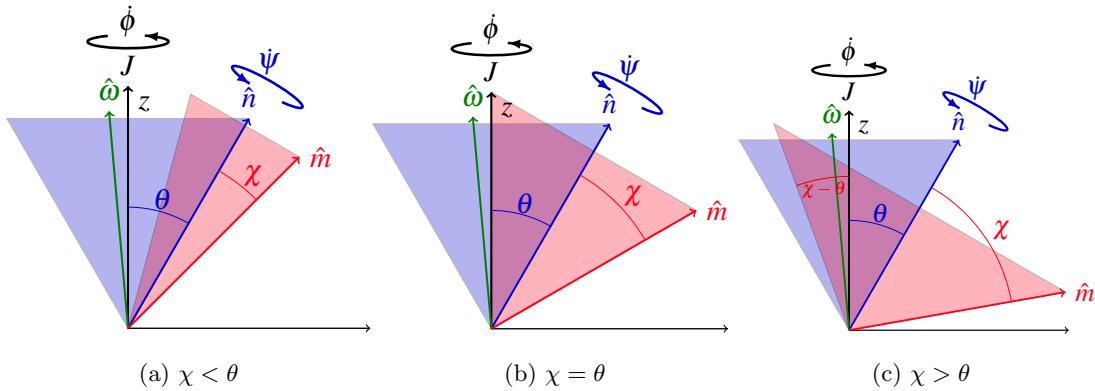


Figure 4.7: Diagrams depicting the projections of the cone swept out by \hat{m}_a about \hat{n}_a (in red) and the cone swept out by \hat{n}_a about J_a (in blue) onto the reference plane for torque free precession. We show the three orderings of χ and θ .

cone, at $\theta + \chi$ the precession cone motion acts in the opposing direction to the spin cone. This causes a reduction in the spin frequency away from the average. By contrast, at $\theta - \chi$ the precession cone motion acts in the same direction as the spin cone producing an increase in the spin frequency above the average.

- The $\chi = \theta$ case: in this special case the angular momentum vector sits exactly on the side of the precession cone. This suggests that at certain precessional phases \hat{m}_a can align exactly with the angular momentum. When this happens the spin frequency tends to zero manifesting as sharp dips in the spin frequency; at the same time the polar angle tends to zero.
- The $\chi > \theta$ case ($\chi = 4\theta$): The precession cone now extends over the angular momentum vector. This means it always acts to reduce the spin frequency; as a result the spin frequency has an average value of $\dot{\phi} + \dot{\psi}$. The polar angle can vary between $\theta + \chi$ and $\chi - \theta$. For χ close to θ the deviations away from the average are large while as χ increases the deviations get smaller as the half angle of the dipole cone increases.

4.3.4 The effective wobble angle

The angle θ (see Figure 4.4) is referred to as the *wobble angle* since, in torque-free precession, its magnitude determines the ‘amount’ of precession. Taking Eqn. (4.33) at an instant when $\Omega_y = 0$ without loss of generality it can be shown that

$$\tan \theta = \frac{I_0}{I_0 + \Delta I} \tan(a) = \frac{1}{1 + \epsilon_I} \tan(\theta + \hat{\theta}). \quad (4.45)$$

Using the angle sum identity for $\tan(\theta + \hat{\theta})$ and expanding about $\epsilon_I = 0$ it can be shown [90] that

$$\hat{\theta} \approx \epsilon_I \sin \theta \cos \theta. \quad (4.46)$$

Then for $\theta \ll 1$, we have that $a \approx \theta + \mathcal{O}(\epsilon_I)$.

When including the EM torque the wobble angle does not have the intuitive interpretation of the ‘amount’ of precession. In particular, the anomalous part of the torque defined in Eqn. (3.5) can be understood to create an effective rotating frame as shown in Section 3.4.1. This means that solutions exist which appear to undergo ‘persistent precession’ where $\theta \approx a \neq 0$ such that the body remains misaligned from the principal axes of its moment of inertia tensor. However, we demonstrated that in fact the body has aligned with the principal axes of its *effective* moment of inertia tensor.

Let us then define an effective wobble angle

$$\tilde{\theta} = \theta - \beta, \quad (4.47)$$

which describes the amount of precession under the full electromagnetic torque. That is, we rotate the usual wobble angle by β defined in Eqn. (3.37). In the limit where $\beta \rightarrow 0$ we recover the usual wobble angle referred to by Jones and Andersson [90] $\theta = \tilde{\theta}$. But, if the effects of the anomalous torque are important (see Chapter 3 for examples of when this is the case), then the effective wobble angle $\tilde{\theta}$ will measure the amount of precession (i.e. its magnitude determines the amplitude of precession). This definition will be used later on when setting up solutions which initially have a minimal amount of precession.

4.4 Observable features: the phase residual

The principal observational quantity reported on for a pulsar is the timing residual. This is the difference between the measured TOA of a pulse and a timing model of the pulsar as discussed in Section 1.2. The phase modulation of pulsars due to precession was first modelled in response to periodic variations observed in the frequency of the Crab pulsar [149, 41]. Precession as a candidate for periodic variations in Hercules X-1 was studied by Bisnovatyi-Kogan et al. [30] and Bisnovatyi-Kogan and Kahabka [31]. Since then, there has been extensive work in the literature: Nelson et al. [128] calculated phase residuals both analytically and numerically for freely precessing stars; torqued precession was considered for a general torque by Jones [93] and Cordes [44] and then for the Deutsch [54] torque by Melatos [121, 122]. In this section, we will calculate residuals by calculating the pulsar phase from our numerical model and then fitting and removing a Taylor expansion. This means that we capture the data collection methods present in any real data. Note that timing residuals and phase residuals are equivalent up to factors of the pulse period. We will discuss this once we have defined the phase residual.

The azimuthal angle Φ given in Eqn. (4.42) gives us the phase of the magnetic dipole. For the observer, the pulsation occurs when the dipole passes through the plane containing the observer and the angular momentum vector. We can always reorient the observer such that the pulsation occurs every time Φ is a multiple of 2π . In this way, Φ is also the phase of the observed pulsations. To generate Φ we define the star’s properties i.e. the

magnetic field and initial angles, then numerically evolve Eqn. (4.7) to Eqn. (4.12). The resulting time series of the spin-vector components and Euler angles are then substituted into Eqn. (4.42) to generate the exact phase.

Following the methods used by observers, we define our ‘timing model’ as a Taylor expansion of the phase about some reference time t_{ref} up to \ddot{f} :

$$\Phi_{\text{fit}}(t; t_{\text{ref}}, f, \dot{f}, \ddot{f}) = 2\pi \left(f(t - t_{\text{ref}}) + \frac{\dot{f}}{2!}(t - t_{\text{ref}})^2 + \frac{\ddot{f}}{3!}(t - t_{\text{ref}})^3 \right), \quad (4.48)$$

with the timing properties (f, \dot{f}, \ddot{f}) as free parameters. Higher order terms can be included, as discussed in Section 2.2, but here we truncate at \ddot{f} . We do not include an initial phase since we can arbitrarily define our reference time to coincide with a pulsation such that the initial phase is zero. Unlike pulsar astronomers, we do not need to worry about other corrections such as the motion of the Earth since Φ is given in the inertial frame.

The phase residual is then the difference between the exact phase and the fitted phase

$$\Delta\Phi(t) = \Phi(t) - \Phi_{\text{fit}}(t; t_{\text{ref}}, f, \dot{f}, \ddot{f}). \quad (4.49)$$

We then use a least-squares fitting method to minimise the root mean squared error of the phase residual. The fitted coefficients $\{f, \dot{f}, \ddot{f}\}$ constitute our best-fit phase model: the best fit to the data $\Phi(t)$ of a power law spin-down model. It is worth noting that a residual depends on which section of data was used in the fit.

In this work, we will report only on phase residuals. However, we note that the phase residual can also be re-scaled to give the timing residual by calculating the residual as a fraction of a cycle then multiplying by the period:

$$\Delta T = \frac{\Delta\Phi(t)}{2\pi} P. \quad (4.50)$$

In the following subsections, we will compare analytic results for the phase residuals due to precession with our numerical results found by fitting and removing a polynomial. As shown in Section 4.3.3 the form of solution will depend on whether $\chi < \theta$ or $\chi > \theta$. Of these two choices, Jones and Andersson [90] concluded that ‘the wobble angle of rapidly rotating stars are limited to small values for the finite crustal breaking strain’; following this reasoning we will limit our study to the $\chi > \theta$ case although the numerical code is capable of finding solutions for either.

4.4.1 Effect of precession on the phase residual

The effect of free precession on phase residuals was first considered by Nelson et al. [128]; since the modulations are due to the geometric effect of precession, they are referred to as the *geometric effect*. To calculate a phase residual, Jones and Andersson [90] subtracted the secular phase evolution from Eqn. (4.43) and, found that the phase residual when

$\chi > \theta$ is given by

$$\Delta\Phi_{49}(t) = -\theta \cot \chi \cos \left(\dot{\psi} t + \frac{\pi}{2} \right), \quad (4.51)$$

where the subscript here refers to the equation number of Jones and Andersson [90]. For a precessing star $\dot{\psi}$ is given by Eqn. (4.38) and the initial value of the cosine argument is $\psi(t=0) = \pi/2$ as derived in Section 4.1.3.

Equation (4.51) is calculated in the absence of any EM torque. Nevertheless, it is still appropriate when a torque is applied provided that the geometric effect is stronger than any other (this is discussed in the next few sections). As such, we begin by simulating a star with the properties listed in Table 4.1. This uses typical values for the angular spin frequency, magnetic field, and a small, but non-zero angle a . The resulting phase residual, in cycles, is given in Figure 4.8. This figure shows strong agreement between the numerical result and the analytic prediction of Eqn. (4.51).

Simulation parameters	
ω_0	= 10 rad/s
B_0	= 10^{13} G
χ	= 50°
a_0	= 2°
\mathcal{A}_{EM}	= 4.3×10^{-6}

Table 4.1: Simulation parameters used for the phase residual plotted in Figure 4.8. Note that \mathcal{A}_{EM} is the electromagnetic amplification factor defined later in Eqn. (4.54).

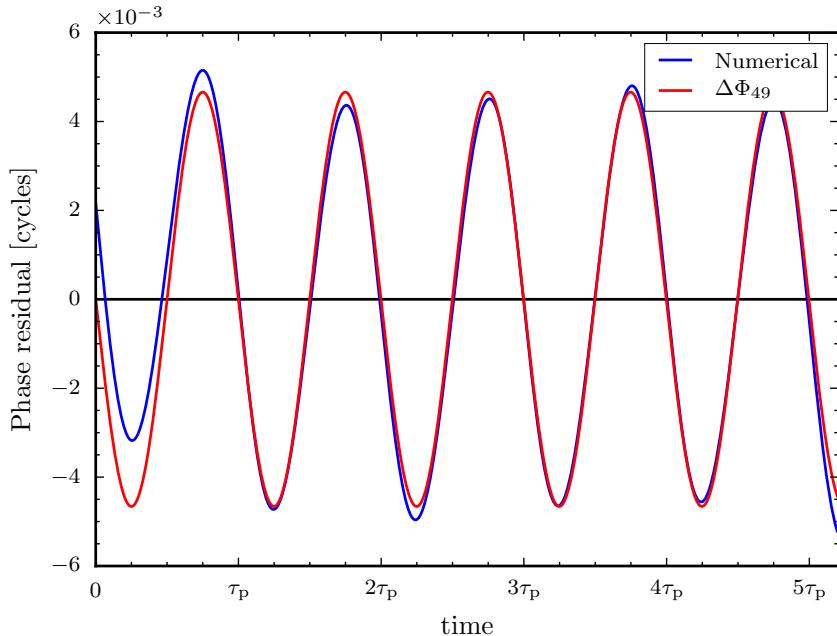


Figure 4.8: The simulated phase residual in cycles for a simulated star with the properties described in Table 4.1. We also plot the corresponding analytic prediction of Eqn. (4.51) which is the geometric effect of precession on the phase residual.

4.4.2 Effect of torqued precession on the phase residual: electromagnetic amplification

The geometric effects of precession can be amplified by the EM torque [44]. Using a simple description based on a vacuum point-dipole spin-down torque, and calculating the departure from a non-precessing power-law spin-down, the amplified phase residuals are given by

$$\Delta\Phi_{63} = -\frac{1}{\pi} \left(\frac{\tau_P}{P} \right) \left(\frac{\tau_P}{\tau_{\text{age}}} \right) \theta \cot \chi \sin(\dot{\psi}t + \pi/2), \quad (4.52)$$

where $\tau_{\text{age}} = |\dot{\Phi}/\ddot{\Phi}|$, note this is equivalent up to a factor 2 with the original definition in Eqn. (1.10); here, we will use the definition from Jones and Andersson [90] for consistency. This result is derived in Jones and Andersson [90] for the magnitude of modulations. Here we keep the exact time evolution behaviour intact. We will derive this expression later on in Section 4.5.1. Notably, we can write the magnitude of modulations in term of Eqn. (4.51) as

$$|\Delta\Phi_{63}| = \frac{1}{\pi} \left(\frac{\tau_P}{P} \right) \left(\frac{\tau_P}{\tau_{\text{age}}} \right) |\Delta\Phi_{49}| \quad (4.53)$$

The two ratios of timescales define an ‘amplification factor’

$$\mathcal{A}_{\text{EM}} = \left(\frac{\tau_P}{P} \right) \left(\frac{\tau_P}{\tau_{\text{age}}} \right). \quad (4.54)$$

The amplification, as first noted by Cordes [44], increases the magnitude of phase residuals for young pulsars with short periods. It is worth noting however, that there is a relative phase difference between Eqn. (4.51) and Eqn. (4.52).

To verify this amplification, we simulate a star using the properties in Table 4.2 for which we use a larger angular spin frequency (this aids in speeding up the computation) and a larger magnetic field than the properties listed in Table 4.1: this results in an amplification factor which is greater than unity. The resulting phase residual is plotted in Figure 4.9 along with the predictions of Eqn. (4.52) and Eqn. (4.51). This clearly demonstrates that the amplified residuals agree with the results calculated numerically while the geometric effect is both out of phase and has a smaller amplitude by a factor $\mathcal{A}_{\text{EM}}/\pi$, where the amplification factor is given in Table 4.2.

Simulation parameters	
ω_0	$= 10^4 \text{ rad/s}$
B_0	$= 10^{14} \text{ G}$
χ	$= 50^\circ$
a_0	$= 2^\circ$
\mathcal{A}_{EM}	$= 43$

Table 4.2: Simulation parameters used for the phase residual plotted in Figure 4.9 and the spin-down rate plotted in Figure 4.10.

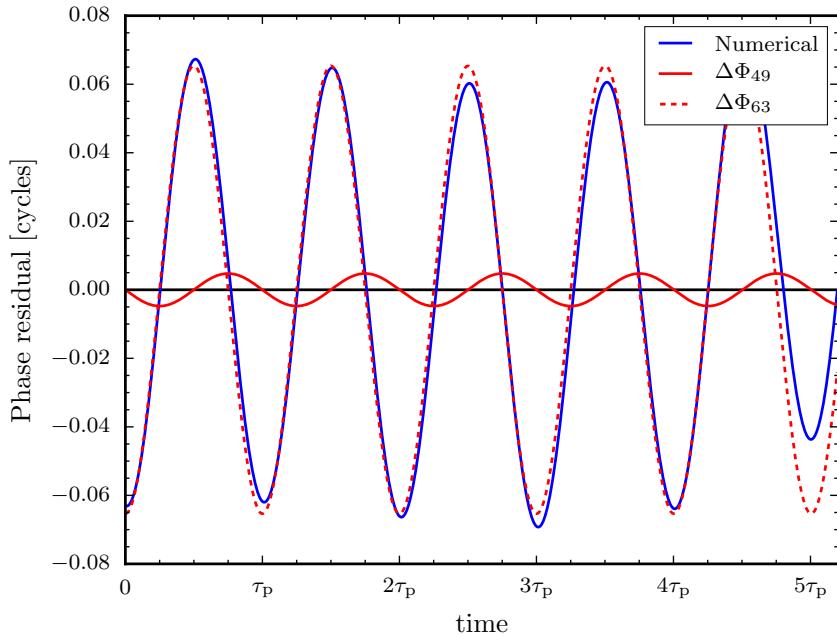


Figure 4.9: The simulated phase residual in cycles for a simulated star with the properties described in Table 4.2. We also plot the corresponding analytic prediction for the geometric effect of precession given by Eqn. (4.51) and the EM amplification of precession given by Eqn. (4.52) which is the appropriate prediction for this simulation.

4.5 Observable features: the spin-down rate

The derivatives of the phase (the frequency, spin-down rate, etc.) display similar periodic departures from their secular evolutions; in the same way as for the phase residual, these periodic modulations can be amplified by the EM torque. In Figure 4.6 we have already shown the modulations of the frequency due to free precession. However, long term modulations of the frequency (on top of the usual secular spin-down), are not reported in the literature and so we will not consider them further. The second observable feature that we will consider is the spin-down rate for which modulations are reported in the literature.

PSR B1828-11 is a normal radio pulsar which displays strong periodic modulation in its spin-down rate. This has been modelled analytically by several authors [156, 90, 108, 14] who all found that the timing residuals and spin-down rate modulations were consistent with EM-dominated precession. We can confirm this by noting that the inferred precession period for this pulsar is $\tau_p \approx 500$ days, the period is $P \approx 0.405$ s, and the spin-down age is $\tau_{\text{age}} \approx P/\dot{P} \approx 2.1 \times 10^5$ yrs (values taken from the ATNF catalogue Manchester et al. [117]), therefore the EM amplification factor defined in Eqn. (4.54) is

$$\mathcal{A}_{\text{EM}} \approx 700. \quad (4.55)$$

In this section, we will derive an analytic model for the spin-down rate of a precessing pulsar dominated by the EM torque; this model will be used in Chapter 5 to test the precession model for PSR B1828-11. Then, we will compare this analytic model with a spin-down rate calculated numerically. We will not consider the geometrically dominated spin-down rate modulations. Further details on that subject can be found in Jones and Andersson [90].

4.5.1 Derivation of the precession spin-down rate

Let us now derive the spin-down rate for a precessing pulsar under a vacuum point-dipole spin-down torque. Equivalent derivations can be found in Link and Epstein [108] and Jones and Andersson [90]; in this derivation we will start with a generalisation of vacuum point-dipole torque to allow for a braking index $n \neq 3$, but retain the angular dependence. The following provides the complete details of a calculation we will use later in Chapter 5 and which was published in Ashton et al. [25].

Our generalisation of the vacuum point-dipole spin-down torque can be written as

$$\ddot{\Phi} = -k\dot{\Phi}^n \sin^2 \Theta, \quad (4.56)$$

where k is a positive constant. From Eqn. (4.43) we have that

$$\cos \Theta = \sin \theta \sin \psi(t) \sin \chi + \cos \theta \cos \chi. \quad (4.57)$$

Rearranging and expanding about $\theta = 0$ and keeping terms up to $\mathcal{O}(\theta^2)$, we find

$$\sin^2 \Theta = \left[\sin^2 \chi + \theta^2 \left(\cos^2 \chi - \frac{\sin^2 \chi}{2} \right) \right] - 2\theta \sin \chi \cos \chi \sin \psi(t) + \frac{1}{2}\theta^2 \sin^2 \chi \cos(2\psi(t)), \quad (4.58)$$

where the first term, in square brackets, is a constant while the second two terms provide the first and second harmonic modulations in $\sin^2 \Theta$.

In order to find approximate solutions to Eqn. (4.56), we begin by defining the time-averaged (constant) value of $\sin^2 \Theta$ as

$$\sin^2 \Theta_0 = \sin^2 \chi + \theta^2 \left(\cos^2 \chi - \frac{\sin^2 \chi}{2} \right), \quad (4.59)$$

then solving Eqn. (4.56) with this constant value we get

$$\dot{\Phi}(t) = \left[\frac{1}{\dot{\Phi}_0^{(n-1)}} + (n-1)k \sin^2 \Theta_0 t \right]^{\frac{-1}{n-1}}, \quad (4.60)$$

where $\dot{\Phi}_0 = \dot{\Phi}(0)$. Now the spin-down timescale is

$$\tau_{\text{age}} = \frac{|\dot{\Phi}_0|}{|\ddot{\Phi}_0|} \approx \frac{1}{k|\dot{\Phi}_0|^{(n-1)} \sin^2 \Theta_0}, \quad (4.61)$$

again we note that this differs by a factor of 2 to the definition in Eqn. (1.10), we use this definition here for consistency with Jones and Andersson [90]. Rearranging gives

$$k = \frac{1}{\tau_{\text{age}} |\dot{\Phi}_0|^{(n-1)} \sin^2 \Theta_0}. \quad (4.62)$$

Then Eqn. (4.60), our zeroth-order solution to the spin-down (under a constant Θ), can be written as

$$\dot{\Phi}(t) = \dot{\Phi}_0 \left[1 + (n-1) \frac{t}{\tau_{\text{age}}} \right]^{\frac{-1}{n-1}}. \quad (4.63)$$

Now, we substitute Eqn. (4.63) back into Eqn. (4.56) along with the expanded, but complete variation in $\sin^2 \Theta$, given by Eqn. (4.58), giving

$$\begin{aligned} \ddot{\Phi}(t) &= \frac{-1}{\tau_{\text{age}} |\dot{\Phi}_0|^{(n-1)} \sin^2 \Theta_0} \dot{\Phi}_0^n \left[1 + (n-1) \frac{t}{\tau_{\text{age}}} \right]^{\frac{-n}{n-1}} \\ &\times \left(\sin^2 \Theta_0 - 2\theta \sin \chi \cos \chi \sin \psi(t) + \frac{1}{2} \theta^2 \sin^2 \chi \cos [2\psi(t)] \right). \end{aligned} \quad (4.64)$$

In principle, this is complete and constitutes an approximation of the spin-down under precession. To simplify this expression, we expand the second bracket with $t/\tau_{\text{age}} \ll 1$, then

$$\begin{aligned} \ddot{\Phi}(t) &= \frac{1}{\tau_{\text{age}}} \dot{\Phi}_0 \left(-1 + n \frac{t}{\tau_{\text{age}}} + \frac{1}{\sin^2 \Theta_0} \left(2\theta \sin \chi \cos \chi \sin \psi(t) - \frac{1}{2} \theta^2 \sin^2 \chi \cos(2\psi(t)) \right) \right) \\ &+ \mathcal{O} \left(\left(\frac{t}{\tau_{\text{age}}} \right)^2, \theta \left(\frac{t}{\tau_{\text{age}}} \right) \right) \end{aligned} \quad (4.65)$$

Next we expand $1/\sin^2 \Theta_0$, from Eqn. (4.59), in θ :

$$\left(\sin^2 \chi + \theta^2 \left(\cos^2 \chi - \frac{\sin^2 \chi}{2} \right) \right)^{-1} \approx \frac{1}{\sin^2 \chi} + \mathcal{O}(\theta^2). \quad (4.66)$$

Note that we only expand to this order since this term is multiplied by a factor of θ and we are neglecting $\mathcal{O}(\theta^3)$ terms. Then simplifying we have

$$\ddot{\Phi}(t) = \frac{1}{\tau_{\text{age}}} \dot{\Phi}_0 \left(-1 + n \frac{t}{\tau_{\text{age}}} + \left(2\theta \cot \chi \sin \psi(t) - \frac{1}{2} \theta^2 \cos(2\psi(t)) \right) \right). \quad (4.67)$$

Finally, converting to spin-frequencies and periods

$$\dot{\nu}(t) = \frac{1}{P_{\tau_{\text{age}}}} \left(-1 + n \frac{t}{\tau_{\text{age}}} + 2\theta \cot \chi \sin \psi(t) - \frac{1}{2} \theta^2 \cos(2\psi(t)) \right) \quad (4.68)$$

We have not made any assumptions on $\psi(t)$ in this derivation. However, since we are interested in cases where $\tau_P \ll \tau_{\text{age}}$ we can make an assumption that

$$\psi(t) = \dot{\psi}t + \psi_0. \quad (4.69)$$

From Figure 4.3 (where $\tau_P/\tau_{\text{age}} \approx 7 \times 10^{-4}$) we see that the torque causes modulations in $\dot{\psi}$. However, in all the cases where we apply this formulae, these modulations will be negligible.

The $\mathcal{O}(\theta)$ modulation term in this expression can be integrated twice to derive Eqn. (4.52); this reflects that this spin-down rate prediction includes the amplification of the EM torque. When $\mathcal{A}_{\text{EM}} < 1$, Eqn. (4.68) is not suitable since the geometric variations (derivatives of Eqn. (4.51)) will dominate. We will not consider the geometric dominated spin-down rates in this work, although our simulation code was tested against known analytic results and found to agree.

4.5.2 Simulations of the precession spin-down rate

In this subsection, we will verify Eqn. (4.68) against our numerical model. In Chapter 3, we concluded that the anomalous torque, which is not included in the derivation of Eqn. (4.68), was not important for realistic pulsars. By verifying that this equation agrees with numerical solutions (calculated with the anomalous torque) we can confirm this conclusion and build confidence in the analytic solution which will be used later in Chapter 5.

To calculate spin-down rates numerically, we could calculate the pulse phase (Eqn. (4.42)) and then numerically differentiate to get $\ddot{\Phi}$. However, when studying the long-term modulations in the spin-down rate, observers (see for example Lyne et al. [111], Perera et al. [134]) use what we will refer to as the ‘observer-method’: a second order Taylor expansion is fit to short sections of data of length T and the resulting coefficient $\dot{\nu}$ is recorded at the mid-point of the section of data; repeating this process every $\sim T/4$ in a ‘sliding-window’ builds a picture of how the spin-down varies with time. We will mimic this data collection process by fitting Taylor expansions to short sections of the simulated phase Φ . This has the benefit that we include any potential peculiarities of the data collection mechanism in our simulation. We choose T such that it is a fraction of the precession period over which we expect quantities to be modulated. This is consistent with the observer-method where T is chosen in order to resolve the observed modulations.

To verify that Eqn. (4.68) accurately models the spin-down rate evolution of precessing pulsars in an EM amplification regime, we simulate a star with the properties given in Table 4.2 (the same set of properties used for the simulated phase residual in Figure 4.9). Most significantly, $\mathcal{A}_{\text{EM}} \gg 1$, which ensures that the EM torque amplifies the spin-down rate modulations. We use an initial spin frequency ω_0 much larger than typical values, this is to allow the numerical solutions to evolve with reasonable computational power; for more physical values the solutions are qualitatively unchanged.

The numerical spin-down rate variations, calculated using the observer-method are plotted along with the analytic predictions of Eqn.(4.68) in Figure 4.10. This figure demonstrates good agreement: small variations result from the application of the observer-method, most notably the numerical results under and over estimates the minimum and maximums due to the time averaging.

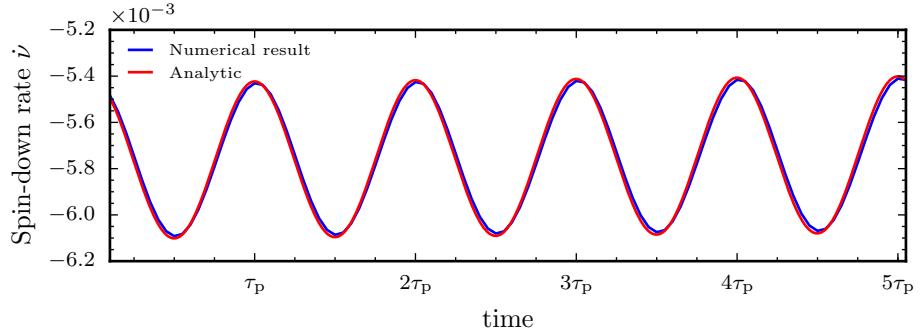


Figure 4.10: Spin-down rate for simulation parameters listed in Table 4.2 compared against the corresponding analytic prediction of Eqn.(4.68).

Figure 4.10 captures the essential features of the spin-down rate of a precessing pulsar amplified by an EM torque. However, there is a special case which produces the ‘double-peaked’ spin-down rate which is observed in PSR B1828-11 [111] (see also Chapter 5) and PSR 0919+06 [134]. This double-peak arises naturally in the precession model when $\theta \ll 1$ and χ is close to $\pi/2$. This can be seen directly from the $\mathcal{O}(\theta^2)$ term in Eqn. (4.68): when $\chi \sim \pi/2$, we get a second harmonic at $\tau_P/2$. To illustrate this, we will repeat the simulation of Figure 4.10, but set $\chi = 85^\circ$; the full set of simulation parameters are listed in Table 4.3. The spin-down rate for this ‘almost orthogonal’ dipole simulation is plotted in the top panel of Figure 4.11 and demonstrates the distinctive double peak.

Simulation parameters	
ω_0	= 10^4 rad/s
B_0	= 10^{14} G
χ	= 85°
a_0	= 10°
\mathcal{A}_{EM}	= 75

Table 4.3: Simulation parameters for the spin-down rate plotted in Figure 4.11.

We can understand this double-peaked feature by recalling that the dipole-cone (introduced in Section 4.3.3) is the cone swept out by the magnetic dipole, \hat{m}_a , at the fast spin frequency; the half-angle this cone makes with the angular momentum vector is Θ (see Eqn. (4.43)). In the lower panel of Figure 4.11 we plot this polar angle for the same simulation. Notably the occurrence of the double-peak in the spin-down rate coincides with times when $\Theta > 90^\circ$. What we are seeing is the magnetic dipole entering the lower hemisphere of the star. By considering the form of the Deutsch [54] torque, we see that if the torque is maximal when the dipole is perpendicular to the spin-vector. Since $\theta \ll 1$ and therefore close to the angular momentum vector, we have a maximum in the

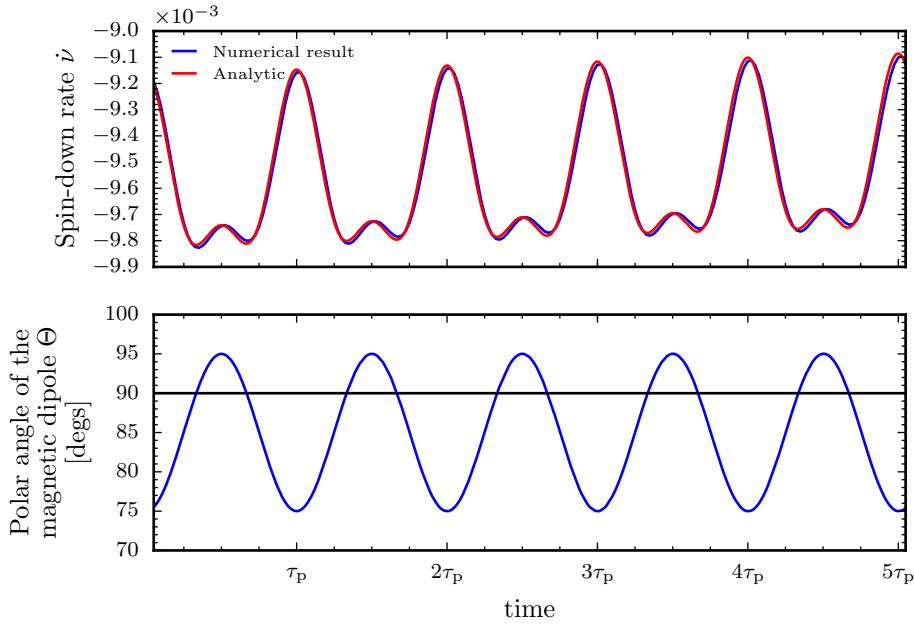


Figure 4.11: Top panel: spin-down rate for simulation parameters listed in Table 4.3 compared against the corresponding analytic prediction of Eqn.(4.68). This simulation differs from Fig. 4.10 in that we chose χ such that the magnetic dipole lies close to the rotation equator. Bottom panel: the variation in the polar angle of the magnetic dipole, Θ , for the simulation.

torque and hence spin-down rate when $\Theta \approx 90^\circ$. As Θ increases past 90° and the dipole enters the lower hemisphere, the absolute value decreases again causing the distinctive double-peak in the spin-down rate.

4.6 Observable features: the pulse profile

So far we have considered observable features which are calculable from the timing properties of the star: the rate at which pulses occur. However, pulsar astronomers also report on the shape of the pulsation by averaging over many pulses to form an integrated pulse profile. Long term variability in pulse profiles is observed in pulsars. The correlated changes in shape for PSR B1828-11 allowed Stairs et al. [156] to rule out a planetary precession hypothesis. More examples of this variability can be found in Lyne et al. [111].

To quantify the changing shape of pulsations from PSR B1828-11, astronomers calculate the *shape parameter* S . This is defined by first noting that PSR B1828-11 alternates between two ‘modes’, one narrow and one broad, then defining S as the fraction of the mean pulse shape attributed to the narrow component (for more details see Stairs et al. [157]). The effect of precession on S for a variety of beam geometries has been considered by Akgün et al. [14]. However, this shape parameter is limiting as it requires there to be two components. A more general method was introduced by Lyne et al. [111] and

defines the *beam-width* as the width of the pulsation at some fraction of the maximum observed intensity.

In the context of precession, the pulse profile is sensitive to both the geometry of the beam itself and the angle made between the beam and the observer. Let us define an observer fixed in the inertial frame such that they maintain a constant angle ι with the angular momentum of the star J_a . For this observer, a pulse can be defined as the moment the dipole cuts through the plane containing them and the angular momentum vector. At this moment, the angle between them and the dipole will be determined by both $\iota \in [0, \pi]$ and $\Theta \in [0, \pi]$. In particular, the angle between the observer and the beam is given by

$$\Delta\Theta = \Theta - \iota. \quad (4.70)$$

Since ι is fixed, variations in $\Delta\Theta$ come solely from variations in the polar angle Θ . We studied these variations in Figure 4.5 and found that Θ has an average value of the larger of θ or χ , then oscillates about this on the precession period with a magnitude given by the smaller value of θ or χ .

In this section, we aim to numerically simulate variations in the beam-width due to precession. To do this, we first model variations in the pulse intensity and then show how to model the determination of the beam-width. The process described here, will be developed later in Chapter 5 to build a predictive analytic model for the beam-width of PSR B1828-11.

4.6.1 Variations in the pulse intensity

The intensity of radiation received by an observer will depend on the orientation of the magnetic dipole with respect to the observer and the beam geometry. It may presumably be maximal when pointing directly at the observer and fall off as the angle between the two grows. For each rotation of the star, the intensity will peak when the beam cuts the plane containing the observer and the angular momentum vector; at this instant the angle between the observer and the beam is given by Eqn. (4.70). If the star is precessing, then the periodic pulses of intensity due to the azimuthal rotation of the star will be modulated by the slower variations in Θ seen in Figure 4.5.

To model this, we take an observer's azimuth and polar angle to be $(\Phi_{\text{obs}}, \iota)$ and then assume the beam geometry follows a Gaussian profile with a single conal emission. This could later be adapted, for example to include conal emission as modelled by Akgün et al. [14]. For such a model of the beam geometry, the intensity of pulsations will vary with the angular separation of the vector from the centre of the star to the observer and the magnetic dipole vector. By considering the intersection of these vectors with the unit sphere, the angular separation can be shown to be

$$\Delta d = \cos^{-1}(\cos(\Theta) \cos(\iota) + \sin(\Theta) \sin(\iota) \cos(\Phi - \Phi_{\text{obs}})) \quad (4.71)$$

Then, taking a Gaussian beam geometry, the intensity of the pulse will be given by

$$\mathcal{I}(\Theta, \Phi, \iota, \Phi_{\text{obs}}, \rho) = \mathcal{I}_0 \exp\left(-\frac{\Delta d^2(\Theta, \Phi, \iota, \Phi_{\text{obs}})}{2\rho^2}\right) \quad (4.72)$$

where \mathcal{I}_0 is the maximum intensity and ρ is a measure of the width of the Gaussian intensity.

Given a value for \mathcal{I}_0 and ρ , we can use a numerical solution to the governing ODEs to simulate this pulse intensity exactly. This is done in Figure 4.12 for a system where the variations in Θ occur on a timescale not much longer than the pulse period. This nonphysical simulation is intended to show the modulation of the individual pulsations due to precession. We can also predict the maximum pulse intensity at any given instant

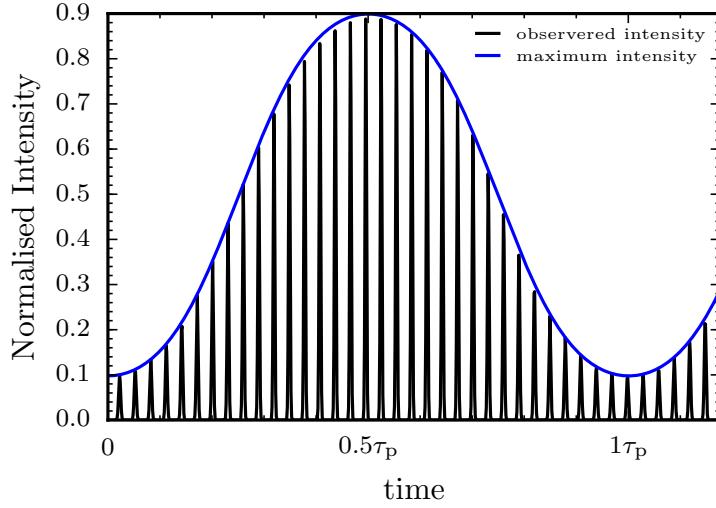


Figure 4.12: Amplitude variation using a 2D Gaussian emission.

by setting $\Phi = \Phi_{\text{obs}}$. Then simplifying, we find that

$$\mathcal{I}_{\text{max}}(\Theta, \Phi, \iota, \rho) = \mathcal{I}_0 \exp\left(-\frac{-(\Theta - \iota)^2}{2\rho^2}\right), \quad (4.73)$$

this is also shown in Figure 4.12.

4.6.2 Variations in the beam-width

It is unlikely that the absolute variations in intensity seen in Figure 4.12 will ever be unambiguously visible in nature. This is because in real observations the intensity will also be subject to variations in the amount of scintillation from the interstellar medium. Therefore, pulsar astronomers do not typically report on intensities themselves, but characterise the pulsation by their beam-width. This is the width of the pulse at some percentage p of the observed maximum intensity. Note that this is not the maximum intensity that the beam produces, \mathcal{I}_0 , but the maximum at that instant in time \mathcal{I}_{max} , as given by Eqn. (4.73)

To calculate the beam-width, we first note that in Eqn. (4.72), Θ varies on the slow precession timescale, while Φ varies on the rapid spin timescale: we are looking to measure the variations with respect to the slow precession timescale. The pulse width is measured by the time spent above a percentage p of the maximum pulse intensity, this can be defined as an inequality

$$\mathcal{I}(\Phi, \Theta, \iota, \Phi_{\text{obs}}, \rho) > \mathcal{I}_{\max} \frac{p}{100}. \quad (4.74)$$

Substituting in Eqn. (4.72) and Eqn. (4.73) then rearranging yields

$$\cos(\Phi - \Phi_{\text{obs}}) > \frac{\cos\left(\sqrt{-2\rho^2 \ln\left[\frac{p}{100} \frac{\mathcal{I}_{\max}}{\mathcal{I}_0}\right]}\right) - \cos(\Theta) \cos(\iota)}{\sin(\Theta) \sin(\iota)} \quad (4.75)$$

where we note that \mathcal{I}_{\max} is not a constant, but will evolve with the polar angle Θ as given by Eqn. (4.43).

Let's consider a single rotation with the magnetic dipole starting and ending in the antipodal point to the observer's position. During this rotation, $\Phi - \Phi_{\text{obs}}$ increases linearly between $-\pi$ and π , and so the left hand side of the inequality is a simple cosine function as illustrated in Figure 4.13. Since we expect Θ to vary slowly compared to the rotation period we can, over a single pulsation, think of Θ as a constant; then the whole right hand side of Eqn. (4.75) is a constant. In Figure 4.13, we illustrate this constant along with the evolution of the left hand side. Then we also illustrate $\Delta\Phi$, the fraction of the rotation period for which the cosine is less than the constant, i.e. Eqn. (4.75) is satisfied.

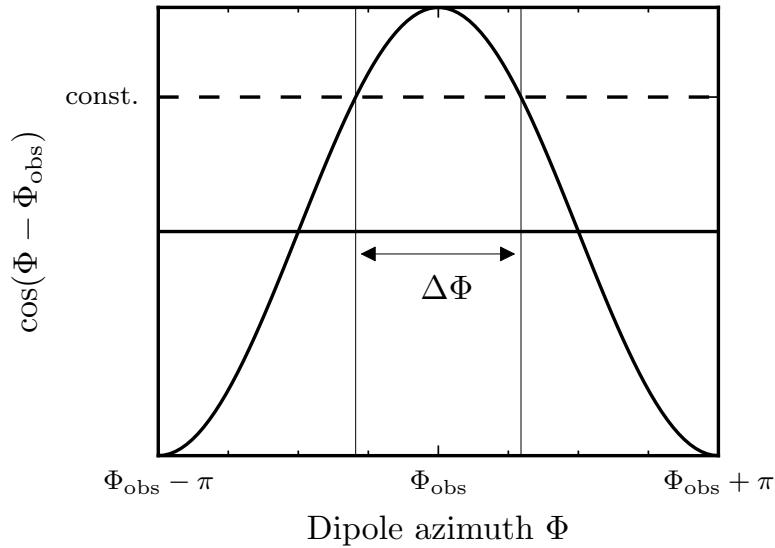


Figure 4.13: Illustration of $\cos(\Phi - \Phi_{\text{obs}})$ during a single rotation. The constant value represents the right hand side of this equation. The width $\Delta\Phi$ indicates the angular period during which Eqn. (4.75) is satisfied.

The angular fraction at which the inequality *is* satisfied is given by $\Delta\Phi$. To convert this into the beam-width reported by pulsar astronomers we multiply by the pulse period. So from $\Delta\Phi$, we can calculate the beam-width as

$$W_p = P \frac{\Delta\Phi}{2\pi} \quad (4.76)$$

where P is the spin period.

The angular width $\Delta\Phi$ when the inequality is satisfied is given by

$$\Delta\Phi = 2 \cos^{-1} \left(\frac{\cos \left(\sqrt{-2\rho^2 \ln \left[\frac{p}{100} \frac{\mathcal{I}_{\max}}{\mathcal{I}_0} \right]} \right) - \cos(\Theta) \cos(\iota)}{\sin(\Theta) \sin(\iota)} \right), \quad (4.77)$$

and so the beam-width is

$$W_p = \frac{1}{\pi \dot{\Phi}} \cos^{-1} \left(\frac{\cos \left(\sqrt{-2\rho^2 \ln \left[\frac{p}{100} \frac{A_{\max}}{A_0} \right]} \right) - \cos(\Theta) \cos(\iota)}{\sin(\Theta) \sin(\iota)} \right). \quad (4.78)$$

Finally, we can also simplify this results using Eqn. (4.73) to give

$$W_p = \frac{1}{\pi \dot{\Phi}} \cos^{-1} \left(\frac{\cos \left(\sqrt{(\Theta - \iota)^2 - 2\rho^2 \ln \left(\frac{p}{100} \right)} \right) - \cos(\Theta) \cos(\iota)}{\sin(\Theta) \sin(\iota)} \right). \quad (4.79)$$

To demonstrate an example of the beam-width modulation, in Figure 4.14 we plot Θ calculated from a numerical simulation with the simulation parameter listed in Table 4.4 and W_{10} as calculated from Eqn. (4.79) with $\iota = 82^\circ$ and $\rho = 0.3$.

Simulation parameters	
ω_0	= 1 rad/s
B_0	= 0 G
χ	= 80°
a_0	= 6°
\mathcal{A}_{EM}	= 0

Table 4.4: Simulation parameters for the beam-width modulations plotted in Figure 4.14.

This shows that there is periodic modulation of the beam-width, with an interesting two-peak structure. This structure can be understood by realising that the polar angle of the dipole Θ ‘passes over’ the observer such that for some of the precessional phase the observer views the beam from below, and other from above. This leads to the two-peaked structure due to the symmetry of the emission geometry. This is a special case which depends on the choices of θ , χ , and ι : more generally the beam-width has a single period oscillatory behaviour.

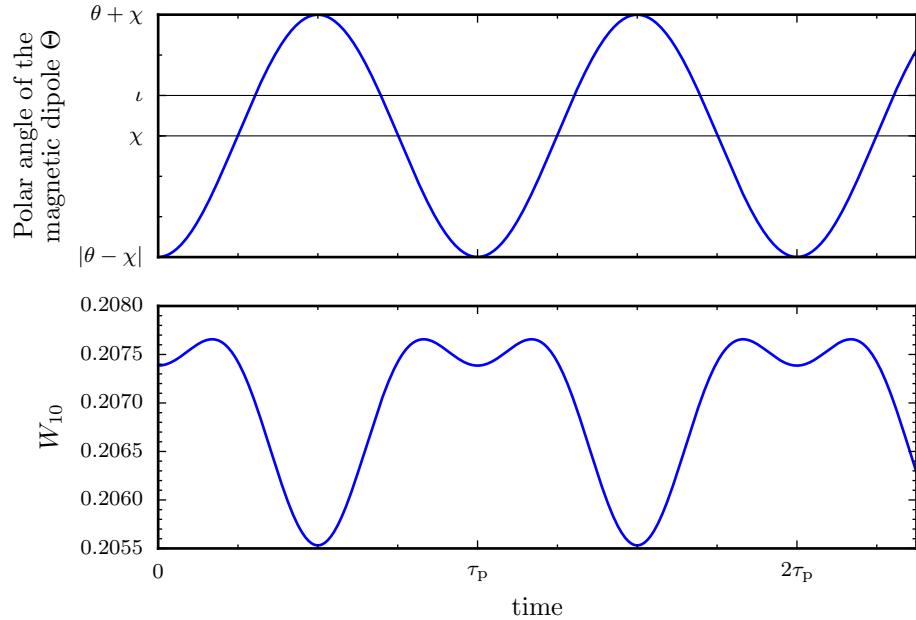


Figure 4.14: Simulation results for the polar angle Θ and W_{10} , a measure of the observed beam-width as calculated from Eqn. (4.79). The simulation parameters for this are listed in Table 4.4 and we additionally set $\iota = 82^\circ$ and $\rho = 0.3$ when calculating W_{10} .

This numerical approach to modelling the beam-width allows us to simulate the effects of precession and probe how this manifests in our observations. We will use the methods discussed in this section later in Chapter 4 when developing and fitting a precession model for the observed beam-width of PSR B1828-11.

4.7 Application: switching and precession

Recently, some workers in the field [111, 134] have suggested that quasi-periodic structure observed in some pulsar timing residuals is a result of magnetospheric torque-switching events as described in Section 2.3.3. In such models, the magnetospheric torque, and hence spin-down rate periodically switches between two distinct values and these changes correlate with changes in the beam-width. These models however are lacking a key feature: the clock which provides the periodicity. It has been suggested [89] that it may in fact be precession which provides this clock. Ultimately, the numerical model developed in this section could study this effect, for example by implementing a hybrid model in which propensity for a magnetospheric switch to occur is related to the angle Θ . In this way, the observed switching would undergo stochastic resonance as suggested by Cordes [45] and discussed in Appendix 2.A.

In this section, we will present novel, but preliminary results on a simplified hybrid magnetospheric switching - precession model. Ultimately we would like to realise the ideas discussed above in a full numerical model. However, for now we will simply consider

the effect of a single magnetospheric torque switching event. Our numerical code is able to model the full hybrid model, but we must first understand the basic physics. In particular, we will simulate a star in which at $t_{\text{switch}} = T_{\text{obs}}/2$ the EM torque suddenly decreases. This is a qualitatively new idea which has not yet been discussed in the literature.

In the EM dipole spin-down model, the torque has two distinct components: the regular spin-down component and the anomalous component. This latter term does not contribute to the spin-down, but as discussed in Section 3.4.1 will modify the axis of precession. Magnetospheric switching models are based on evidence that the spin-down rate is switching, therefore the torque switching must occur in the spin-down component. However, it is unclear if it will also occur in the anomalous component. To answer this, one would need a detailed model of how the magnetosphere reconfigures during the switching. We will not do this here, but instead investigate a simple phenomenological switching event in which we modify Eqn. (3.5) as follows

$$\mathbf{T} = (1 - S_{\text{S}}H(t - t_{\text{switch}}))\mathbf{T}_{\text{S}} + (1 - S_{\text{A}}H(t - t_{\text{switch}}))\mathbf{T}_{\text{A}} \quad (4.80)$$

where the subscripts label the spin-down and anomalous components, S is the strength of switching, and $H(t)$ is the Heaviside step function. In this model we can control which components are switched by choosing S_{S} and S_{A} appropriately.

In our model the strength of the EM torque is parameterised by ϵ_{A} , related to the surface magnetic field strength by

$$\epsilon_{\text{A}} = \frac{R^5}{4I_0c^2}B_0^2. \quad (4.81)$$

Rearranging Eqn. (1.12) we can then write the spin-down rate as

$$\begin{aligned} \dot{\omega}_0 &= -\frac{B_0^2 R^6 \sin^2(\alpha) \omega_0^3}{6I_0 c^3} \\ &= -\frac{2R\epsilon_{\text{A}} \sin^2(\alpha) \omega_0^3}{3c}, \end{aligned} \quad (4.82)$$

where α is the angle between the spin-vector and magnetic dipole. Since we expect these to be misaligned in order to observe pulsations, we can take $\sin^2 \alpha \approx 1$, then our spin-down rate is approximately

$$\dot{\nu} = -\frac{R\omega_0^3}{3\pi c}\epsilon_{\text{A}}. \quad (4.83)$$

From this we can equate the switching in the spin-down rate S_{S} directly to that measured from pulsar observations. That is, from Eqn. (4.80) we have that

$$\epsilon_{\text{A}} \rightarrow \epsilon'_{\text{A}} = (1 - S_{\text{S}})\epsilon_{\text{A}}. \quad (4.84)$$

Then the spin-down rate also changes as

$$\dot{\nu} \rightarrow \dot{\nu}' = (1 - S_S)\dot{\nu}. \quad (4.85)$$

during a switching event.

4.7.0.1 Minimal precession initial state

In the following section we will investigate the effects of a single switch, but first we choose to define an initial ‘minimal precession’ from which to begin our simulations. This is done to mimic the state of real pulsars, which in general are not found to be precessing. Moreover, it helps to discern the effects of the switching from that of precession.

Precession will not occur when the spin-vector is aligned with the principal axis of the effective moment of inertia tensor (defined in Sec. 3.4.1). The angle between these is approximately the effective wobble angle defined in Eqn. (4.47). For minimal precession we should therefore set this wobble angle to zero. In all simulations, we consider a biaxial body with the full torque given by (4.80). From our previous discussion on the wobble angle we can minimise the precession initially by setting the initial polar angle of the spin-vector in the rotating frame to lie along the effective body-frame axis. In the presence of the anomalous torque this is done with

$$a_0 = \beta(\epsilon_I, \epsilon_A, \chi), \quad (4.86)$$

where β is defined in Eqn. (3.37).

In Figure 4.15 we illustrate the phase residual of our simulation in the absence of a switching event; the selected simulation parameters are listed in Table 4.5. We use a large magnetic field so that the effect of switching can be easily identified, a typical angular spin frequency, and an initial polar angle which is exactly the angle β , as calculated using Eqn. (3.37).

Simulation parameters	
ω_0	= 100 rad/s
B_0	= 10^{15} G
χ	= 50°
a_0	= -0.78°
\mathcal{A}_{EM}	= 23

Table 4.5: Simulation properties used for Figure 4.15, Figure 4.16, and Figure 4.17.

Notably this minimal precession solution does show some precession with phase residuals $\sim 10^{-5}$. This is because the spin-down torque produces a wobble in the angular momentum vector and as a result the wobble angle is not truly zero. For all the phase residuals calculated in this section, we will fit and remove a Taylor expansion up to and including $\ddot{\nu}$.

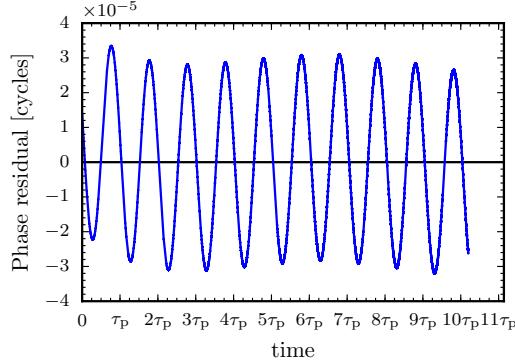


Figure 4.15: The phase residual for a minimal precession simulation with no switching event; the simulation properties are listed in Table 4.5. In this case, we have fitted and removed a Taylor expansion up to and including $\ddot{\nu}$. Note that, due to inherent errors in the fitting process, the residual exhibits an apparent periodicity which matches its span (or some integer number of periods fits into the span, in this case 2). This was shown to be the case by varying the observation span and observing a corresponding shift in the period. As such it is a non-physical effect that can be ignored.

We will now set up simulations of this ‘minimal precession’ star, and then manually switch the torque. We choose a star where the EM torque amplification is important. In the following sections, we will consider first switching in the spin-down torque, and then also in the anomalous torque. For all simulations considered here we will use the properties listed in Table 4.5.

4.7.1 Switching in the spin-down torque only

We now consider manually switching the spin-down torque halfway through the simulation, with no switch occurring in the anomalous component. That is we set

$$t_{\text{switch}} = \frac{T_{\text{obs}}}{2}, \quad S_S = 0.4, \quad S_A = 0.0, \quad (4.87)$$

such that halfway through the simulation the spin-down torque is reduced by a fraction 0.4 while the anomalous torque remains unaffected.

In Figure 4.16 we plot the phase residuals from this simulation. In the top plot is the residual as calculated over the entire observation period. We find a single periodic variation with a period of the observation time T_{obs} : due to sudden change in spin-down rate the Taylor expansion fit is not finding a suitable fit. The precession features which existed in the initial minimal precession are swamped by the larger variations due to the switch. To study this simulation further, in the lower plot we plot two residuals: the first is calculated in the region $[0, t_{\text{switch}}]$ and the second in $[t_{\text{switch}}, T_{\text{obs}}]$. Because the switch does not occur in either of these periods we can resolve the free precession during each period and note that the precession modulation is in fact smaller after the switch. This reduction in the size of modulations is because the precession is due to the

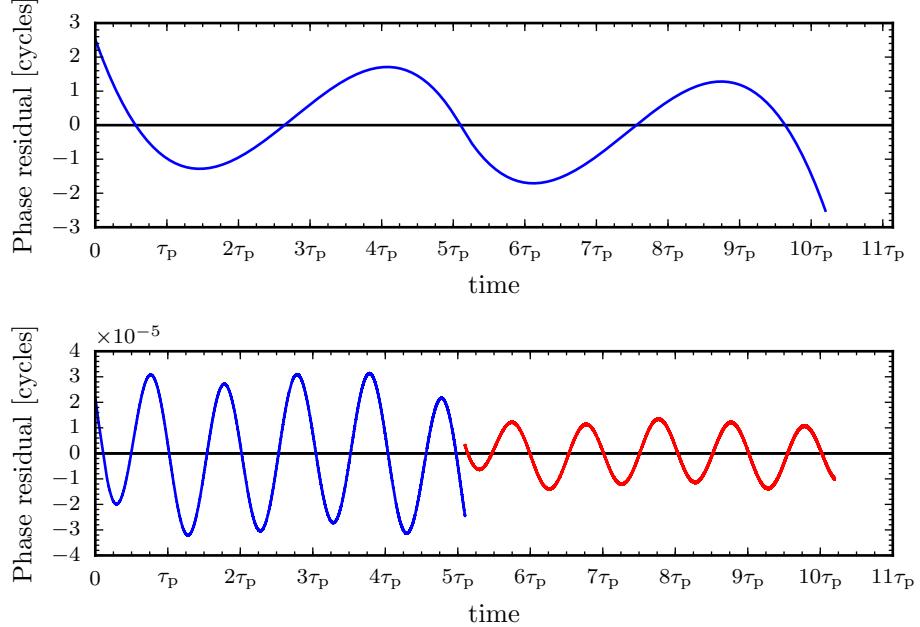


Figure 4.16: Phase residuals for a simulation with a single switch in the spin-down torque with the parameters given in Eqn. (4.87) and the simulation properties listed in Table 4.5. In the top plot we show the residual calculated using the whole observation time. The bottom plot is the separate residuals calculated in the region before and after the switch.

spin-down torque wobbling the angular moment vector and being amplified by the EM torque. After the switch the spin-down torque, and hence \mathcal{A}_{EM} , is reduced by a factor S_S and therefore the size of the modulation is similarly reduced.

4.7.2 Switching in the spin-down and anomalous torque

We now consider manually switching both the spin-down and anomalous torque halfway through the simulation. That is we set

$$t_{\text{switch}} = \frac{T_{\text{obs}}}{2}, \quad S_S = S_A = 0.01 \quad (4.88)$$

In a similar fashion to Figure 4.16, we first show the total residual in the top plot of Figure 4.17, and then the individual residuals in the lower plot. Again, the overall phase residual exhibits periodic modulation with a period of the observation time T_{obs} resulting from the failing of the fitting algorithm to find an appropriate fit. In contrast to the spin-down only switching, the amount of precession when considered before and after the glitch now increases.

To understand this, recall that we begin with a minimal precession state, where $\theta = \beta$ and the precession results from effect of the spin-down torque. After the switch, we have changed the size of the anomalous torque and hence we have modified the effective body

frame and the angle β . This means that after the switch the star is no longer in a minimal precession configuration. This generates a significantly larger wobble angle producing a significant increase in the phase residuals fitted in the post-switch period. The effect is not observable when fitting to the entire simulation period since the switching event remains dominant. In the next section we will calculate this new wobble angle.

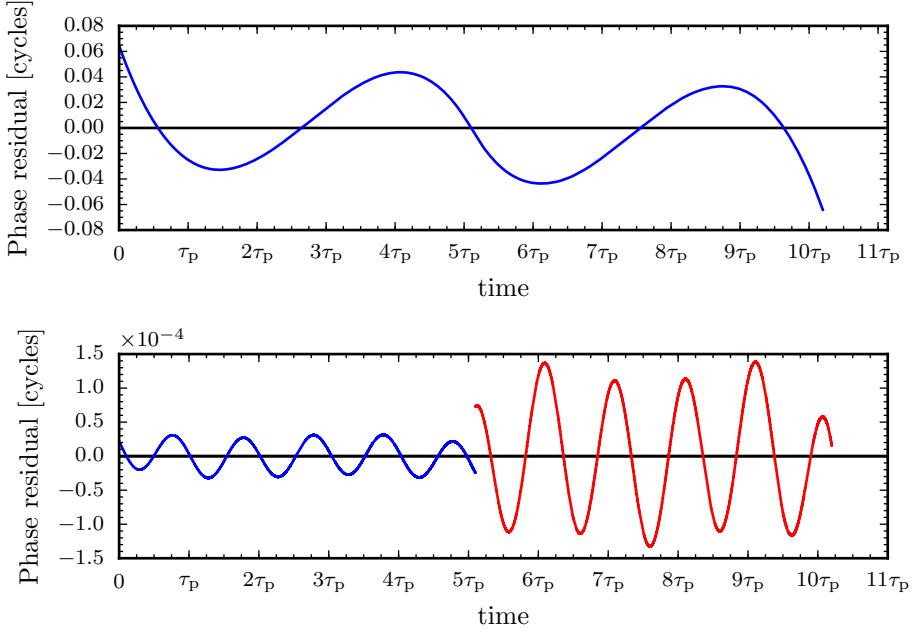


Figure 4.17: Phase residuals for a simulation with a single switch in the spin-down and anomalous torque with the parameters given in Eqn. (4.88) and the simulation parameters listed in Table 4.5. In the top plot we show the residual calculated using the whole observation time. The bottom plot is the separate residuals calculated in the region before and after the switch.

4.7.2.1 Calculating the new wobble angle after a switch

We now estimate the change in wobble angle after switching a fraction of the anomalous torque. This will be useful in making testable predictions for anomalous torque switching.

The two-state switching changes the value of ϵ_A according to Eqn. (4.84), which in turn redefined the effective rotating frame as defined in Section 3.4.1. A non-precessing star at an angle $\beta(\epsilon_I, \epsilon_A, \chi)$ will, after an anomalous torque switch by a fractional amount S_A , no longer be aligned with the rotating frame axis. This is because the effective rotating frame will have shifted to $\beta' = \beta(\epsilon_I, S_A \epsilon_A, \chi)$. As a result, we should expect the previously non-precessing star to begin precessing after a torque switching event. The exact size of the new precession wobble angle will depend on the phase of precession (if it was precessing) at the instant before the switching event. The exact change in the wobble angle will depend on the precessional phase at the time of the glitch. We can

approximate it in a crude way by considering

$$\Delta\beta(\epsilon_I, \epsilon_A, \chi, S_A) \sim |\beta - \beta'|. \quad (4.89)$$

The expression for $\Delta\beta$ is not easily amenable to analytic calculation, but can easily be explored graphically. In Figure 4.18 we plot Eqn. (4.89) by calculating β and β' values using Eqn. (3.37); this is done for several choices of S_A to show the typical variations. This illustrates that the new precession angle after a switch can be as much as a few degrees although it tends to zero in the limit $\epsilon_I \gg \epsilon_A$ and $\epsilon_I \ll \epsilon_A$.

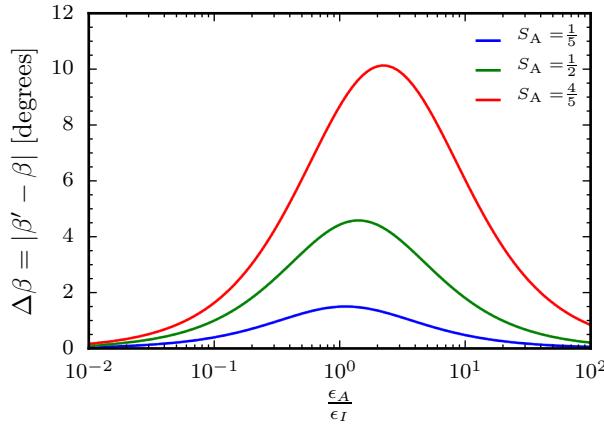


Figure 4.18: Illustrating the magnitude of the precession angle after switching due to the new rotation of the effective rotating frame. We plot the half-angle ($\Delta\beta$) of the precession cone as a function of the ratio ϵ_A/ϵ_I . Typically we expect real stars to have $\epsilon_A < \epsilon_I$.

A simple application of this work is to apply our findings to PSR B1828-11, which was interpreted by Lyne et al. [111] as undergoing switching with a fractional change in the spin-down rate given by

$$\frac{\Delta\dot{\nu}}{\dot{\nu}} = 0.007. \quad (4.90)$$

Manipulating Eqn. (4.85), we see that

$$|S_S| = \frac{\Delta\dot{\nu}}{\dot{\nu}}. \quad (4.91)$$

Using data from the ATNF catalogue [117], PSR B1828-11 has a frequency $\nu = 2.47$ Hz, a spin-down rate $\dot{\nu} = -3.65 \times 10^{-13}$ Hz/s. Rearranging Eqn. (4.83) we then have

$$\epsilon_A^{B1828-11} = \frac{c}{8\pi R\nu^3} \dot{\nu} \approx 3 \times 10^{-11}. \quad (4.92)$$

Now in this interpretation, ϵ_I is unconstrained for PSR B1828-11 since the periodic modulations are assumed to be resulting from switching and not precession. Nevertheless,

we numerically maximise $\Delta\beta$ over ϵ_I to find

$$\Delta\beta_{\max}(\epsilon_A = 2.89 \times 10^{-11}, \epsilon_I = 2.87 \times 10^{-11}) = 0.05 \text{ degs.} \quad (4.93)$$

This is the maximum new wobble angle which must occur every time PSR B1828-11 undergoes a switching event, if the switch also occurs in the anomalous torque. This therefore provides a method to probe how the switching mechanism works by looking at timing residuals in the post-switch timing data to check for any signs of precession.

4.8 Conclusions

In this chapter, we have developed a method to simulate observable properties of neutron stars by evolving the governing system of ODEs. These equations, given in Eqn. (4.7) to Eqn. (4.12), contain both the components of the spin-vector in the rotating frame and the Euler rotation angles required to transform from the rotating frame to the inertial frame. This step is important since it is in this reference frame that observers view the effects of precession on the pulsar's observation features.

We developed an intuitive model for how the magnetic dipole, along which EM radiation is emitted, moves in the inertial frame. Comparing with analytic models where possible, we showed how to calculate from the motion of the dipole the evolution of the star's timing properties (the phase, frequency, and spin-down rate) and how to calculate features of the beam shape such as the beam-width. In doing this, we derived the spin-down rate for a precessing pulsar (see Section 4.5.1), which will be used in Chapter 5, and verified it against numerical solutions.

Finally, we used the phase residuals to investigate the effect of a simple torque switching model in which the components of the torque instantaneously change. This was done to model the magnetospheric switching which is proposed by Lyne et al. [111] as an explanation of the periodic modulation of PSR B1828-11. We present preliminary results showing the complicated interactions which occur in single switching events; in the future we would like to extend this to having the precession determine when the switching occurs. Our numerical model is ideally suited to this task as it captures the complicated feedback between precession and the changing torque. Furthermore, we would eventually like to link the switching to the precessional phase in a probabilistic way as proposed by Jones [89]. In this way, precession provides the clock to the switching. By virtue of being probabilistic, such a system will display stochastic resonance as first described by Cordes [45].

Chapter 5

Comparing models of the periodic variations in spin-down and beam-width for PSR B1828-11

In this thesis, we endeavour to understand the cause of timing variations in pulsars with the aim being that in doing so, we can understand more about the physics of neutron stars. In Chapter 2 we discussed a variety of models to explain timing noise including two, precession and magnetospheric switching, which apply to pulsars with periodic modulations. Such pulsars are particularly interesting given the conclusion of Hobbs et al. [81] that, as more data is collected for pulsars, more quasi-periodic features are observed. Therefore, it is hoped that understanding what causes the periodic modulations may eventually lead to a more complete understanding of timing noise across the population.

In Chapter 3 and Chapter 4, we investigated precession and its effects on the observed features of pulsars. One pulsar, PSR B1828-11, provides strong evidence for precession [156], but has also been put forth as a candidate for magnetospheric switching [111]. There is no decisive feature in the data on PSR B1828-11 which rules out either model. Therefore, in this chapter we will apply a Bayesian model comparison to determine quantitatively, which model best fits the data. In doing this, we will fully develop both the precession model and the switching model into complete predictive models for the spin-down rate and beam-width for which we have observational data. This chapter contains material reprinted with permission from Ashton et al. [25] as follows: Ashton, G.; Jones, D. I.; Prix, R. Volume 458, Issue 1, p.881-899 (2016). Copyright (2016) by the Monthly Notices of the Royal Astronomical Society.

5.1 Introduction

The pulsar PSR B1828-11 demonstrates periodic variability in its pulse timing and beam shape at harmonically related periods of 250, 500, and 1000 days. The modulations in

the timing was first taken as evidence that the pulsar is orbited by a system of planets by Bailes et al. [27]. A more complete analysis by Stairs et al. [156] concluded that the corresponding changes in the beam-shape would require at least two of the planets to interact with the magnetosphere, which does not seem credible. Instead, the authors proposed that the correlation between timing data and beam-shape suggested the pulsar was undergoing free precession. If true, such a claim would require rethinking of the vortex-pinning model used to explain the pulsar glitches since the pinning should lead to much shorter modulation period than observed [152], and fast damping of the modulation [107].

The idea of precession for PSR B1828-11 has been studied extensively in the literature: Jones and Andersson [90] derived the observable modulations due to precession and noted that the electromagnetic spin-down torque will amplify these modulations (see Section 4.4.2 for more details). Link and Epstein [108] fitted a torqued-precession model to the spin-down and beam-shape, followed by Akgün et al. [14] where a variety of shapes and the form of the spin-down torque were tested. All of these authors agree that precession is a credible candidate to explain the observed periodic variations. Furthermore, as we found in Section 4.5.2, to explain the double-peaked spin-down modulations, the so-called wobble angle must be small while the angle between the symmetry axis of the biaxial moment of inertia tensor and the magnetic dipole must be close to $\pi/2$.

More recently Arzamasskiy et al. [22] updated the previous estimates (based on a vacuum approximation) to a plasma filled magnetosphere. They also find that the magnetic dipole and spin-vector must be close to orthogonal, but solutions could exist where it is the wobble angle which is close to $\pi/2$ while the magnetic dipole lies close to the angular momentum vector; we will not consider such a model here, but note it is a valid alternative which deserves testing.

The distinctive spin-down of PSR B1828-11 was analysed by Seymour and Lorimer [151] for evidence of chaotic behaviour. They found evidence that PSR B1828-11 was subject to three dynamic equations with the spin-down rate being one governing variable. This further motivates the precession model since it results from applying Euler's three rigid-body equations to a non-spherical body [100].

The precession hypothesis was challenged by Lyne et al. [111] when reanalysing the data. They noted that measuring the spin-down and beam-shape with any accuracy requires time averaging over periods ~ 100 days, which smooths out any behaviour acting on this time-scale. Motivated by the intermittent pulsar PSR B1931+24, they put forward the phenomenological hypothesis that instead the magnetosphere is undergoing periodic switching between (at least) two metastable states. Such switching would result in correlated changes in the beam-width and spin-down rate. They returned to the data and instead of studying a time-averaged beam-shape-parameter as done by Stairs et al. [156], they instead considered the beam-width at 10% of the observed maximum W_{10} . This quantity is time-averaged, but only for each observation lasting $\sim 1\text{hr}$. This makes W_{10} insensitive to any changes which occur on timescales shorter than an hour. If the meta-stable states last longer than this, W_{10} will be able to resolve the switching. The relevant data was kindly supplied to us courtesy of Andrew Lyne [111], and is

reproduced in Figure 5.1. From these observations, the authors concluded that the individual measurements of W_{10} for PSR B1828-11 did in fact appear to switch between distinct high and low values, as opposed to a smooth modulation between the values, with this switching coinciding with the periodic changes in the spin-down. On this basis, they interpret the modulations of PSR B1828-11 as evidence it is undergoing periodic switching between two magnetospheric states. When studying another pulsar which also displays double-peaked spin-down modulations, Perera et al. [134] extended the switching model, as discussed in Section 5.3.2, to be capable of producing the double-peak spin-down rate; it is this modification of the switching model which we will be comparing with precession.

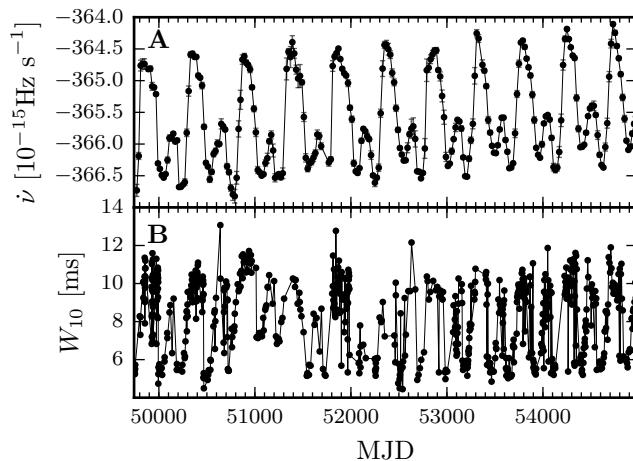


Figure 5.1: Observed data for PSR B1828-11 spanning from MJD 49710 to MJD 54980. In panel **A** we reproduce the spin-down rate with error-bars and in panel **B** the beam-width W_{10} (for which no error bars were available). All data courtesy of Lyne et al. [111].

In our view, it is not immediately clear by eye whether the data presented in Figure 5.1 is sufficient to rule out or even favour either of the precession or switching interpretations. For this reason, we develop a framework in which to evaluate models built from these concepts and argue their merits quantitatively using a Bayesian model comparison. We note that a distinction must be made between a conceptual idea, such as precession, and a particular predictive model built from it. As we will see, each concept can generate multiple models, and furthermore we could imagine using a combination of precession and switching, with the precession acting as the ‘clock’ that modulates the probability of the magnetosphere being in one state or the other, an idea developed by Jones [89]. The models considered here cover the precession and switching interpretations, but we do not claim the models to be the ‘best’ that these hypotheses could produce.

We note that there also exists high-resolution observations of the *shape parameter* of PSR B1828-11 which are presented in Stairs et al. [157] and Lyne [109]. The shape parameter is defined to be the proportion of time spent in one of two modes as determined from the integrated pulse profile. These observations show that PSR B1828-11 appears to ‘switch’ between the modes on short timescales (less than ~ 8 hrs). This, it seems, provides evidence for any model where the switching occurs on short timescales, but is

biased by a longer timescale of ~ 500 days. The effect of precession on the shape parameter was investigated by Akgün et al. [14] who demonstrated how a ‘patchy’ emission region could produce such observations under a pure precession hypothesis. Alternatively, precession could in fact be the clock biasing the switching as in the hybrid model of Jones [89]. For the switching model, this observation also complicates the story since it requires stability over two quite different timescales. In this chapter, we will focus only on the long timescale variations shown in Fig. 5.1 and how these can be interpreted under a deterministic switching model or precession; in effect assuming that the short timescale variations are unrelated to the models under investigation. In the future, we would like to include these data sets and see how this changes the picture.

The rest of the chapter is organised as follows: in Section 5.2 we will describe the framework to fit and evaluate a given model, in Section 5.3 we will define and fit several predictive models from the conceptual ideas, and then in Section 5.4 we shall tabulate the results of the model comparison. Finally, the results are discussed in Section 5.5.

5.2 Bayesian Methodology

We now introduce a general methodology to compare and evaluate models for this form of data. The technique is well practised in this and other fields and so in this section we intend only to give a brief overview; for a more complete introduction to this subject see [155, 86, 64].

5.2.1 The odds-ratio and posterior probabilities

There are two issues that we wish to address. Firstly, given two models, how can one say which is preferred, and by what margin? Secondly, assuming a given model, what can be said of the probability distribution of the parameters that appear in that model?

We can address the first issue by making use of Bayes theorem for the probability of model \mathcal{M}_i given some data:

$$P(\mathcal{M}_i|\text{data}) = P(\text{data}|\mathcal{M}_i) \frac{P(\mathcal{M}_i)}{P(\text{data})}. \quad (5.1)$$

The quantity $P(\text{data}|\mathcal{M}_i)$ is known as the *marginal likelihood* of model \mathcal{M}_i given the data.

In general, we cannot compute the probability given in equation (5.1) because we do not have an exhaustive set of models to calculate $P(\text{data})$. However, we can compare two models, say A and B , by calculation of their *odds-ratio*:

$$\mathcal{O} = \frac{P(\mathcal{M}_A|\text{data})}{P(\mathcal{M}_B|\text{data})} = \frac{P(\text{data}|\mathcal{M}_A)}{P(\text{data}|\mathcal{M}_B)} \frac{P(\mathcal{M}_A)}{P(\mathcal{M}_B)}. \quad (5.2)$$

In the rightmost expression, the first factor is the ratio of the marginal likelihoods (also known as the *Bayes factor*) which we will discuss shortly, while the final factor reflects

our prior belief in the two models. If no strong preference exists for one over the other, we may take a non-informative approach and set this equal to unity. We will follow this approach in what follows below.

We need to find a way of computing the marginal likelihoods, $P(\text{data}|\mathcal{M}_i)$. To this end, consider a single model \mathcal{M}_i with model parameters $\boldsymbol{\theta}$, and define $P(\text{data}|\boldsymbol{\theta}, \mathcal{M}_i)$ as the *likelihood function* and $P(\boldsymbol{\theta}|\mathcal{M}_i)$ as the *prior distribution* for the model parameters. We can then perform the necessary calculations by making use of

$$P(\text{data}|\mathcal{M}_i) = \int P(\text{data}|\boldsymbol{\theta}, \mathcal{M}_i)P(\boldsymbol{\theta}|\mathcal{M}_i)d\boldsymbol{\theta}. \quad (5.3)$$

The likelihood function can also be used to explore the second issue of interest, by calculating the *joint-probability distribution* for the model parameters, also known as the *posterior probability distribution*:

$$P(\boldsymbol{\theta}|\text{data}, \mathcal{M}_i) = \frac{P(\text{data}|\boldsymbol{\theta}, \mathcal{M}_i)P(\boldsymbol{\theta}|\mathcal{M}_i)}{P(\text{data}|\mathcal{M}_i)}. \quad (5.4)$$

Note that the marginal likelihood $P(\text{data}|\mathcal{M}_i)$ described above plays the role of a normalising factor in this equation.

In general, the integrand of Eqn. (5.3) makes analytic, or even simple numeric integration difficult or impossible. This is the case for the probability model that we will use and so instead we must turn to sophisticated numerical methods. For this study we use Markov-Chain Monte-Carlo (MCMC) techniques which simulate the joint-posterior distribution for the model parameters up to the normalising constant

$$P(\boldsymbol{\theta}|\text{data}, \mathcal{M}_i) \propto P(\text{data}|\boldsymbol{\theta}, \mathcal{M}_i)P(\boldsymbol{\theta}|\mathcal{M}_i). \quad (5.5)$$

In particular we will use the Foreman-Mackey et al. [62] implementation of the affine-invariant MCMC sampler [71] to approximate the posterior density of the model parameters. Further details of our MCMC calculations can be found in Appendix 5.A.

Once we are satisfied that we have a good approximation for the joint-posterior density of the model parameters we discuss how to recover the normalising constant to calculate the odds-ratio in Section 5.4.

5.2.2 Signals in noise

We now need to build a statistical model to relate physical models for the spin-down and beam-width to the data observed in Figure 5.1. To do this we will turn to a method widely used to search for deterministic signals in noise.

We assume our observed data y^{obs} is a sum of a stationary zero-mean Gaussian noise process $n(t, \sigma)$ (here σ is the standard deviation of the noise process) and a signal model

$f(t|\mathcal{M}_j, \boldsymbol{\theta})$ (where $\boldsymbol{\theta}$ is a vector of the model parameters) such that

$$y_i^{\text{obs}}(t_i|\mathcal{M}_j, \boldsymbol{\theta}, \sigma) = f(t_i|\mathcal{M}_j, \boldsymbol{\theta}) + n(t_i, \sigma). \quad (5.6)$$

Given a particular signal model, subtracting the model from the data should, if the model and model parameters are correct, leave behind a Gaussian distributed residual - the noise. That is

$$y_i^{\text{obs}}(t_i|\mathcal{M}_j, \boldsymbol{\theta}, \sigma) - f(t_i|\mathcal{M}_j, \boldsymbol{\theta}) \sim N(0, \sigma). \quad (5.7)$$

The data, for either the spin-down or beam-width, consists of N observations (y_i^{obs}, t_i) . For a single one of these observations, the probability distribution given the model and model parameters is

$$P(y_i^{\text{obs}}|\mathcal{M}_j, \boldsymbol{\theta}, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(f(t_i|\mathcal{M}_j, \boldsymbol{\theta}) - y_i)^2}{2\sigma^2}\right\}. \quad (5.8)$$

The likelihood is the product of the N probabilities

$$P(\mathbf{y}^{\text{obs}}|\mathcal{M}_j, \boldsymbol{\theta}, \sigma) = \prod_{i=1}^N P(y_i^{\text{obs}}|\mathcal{M}_j, \boldsymbol{\theta}, \sigma), \quad (5.9)$$

where \mathbf{y}^{obs} denotes the vector of all the observed data.

In Section 5.3 we will define the physical models, $f(t|\mathcal{M}_j, \boldsymbol{\theta})$, for the precession and switching interpretations; for now we recognise that once defined, we may calculate the likelihood of the data under the model using Eqn. (5.9).

5.2.3 Choosing prior distributions

In the previous section we have developed the likelihood function $P(\text{data}|\boldsymbol{\theta}, \mathcal{M}_j)$ for any arbitrary model producing a deterministic signal $f(t_i|\mathcal{M}_j, \boldsymbol{\theta})$ in noise. To compare between particular models, using Eqn. (5.2), we must compute the marginal likelihood as defined in Eqn. (5.3) which requires a prior distribution $P(\boldsymbol{\theta}|\mathcal{M}_j)$.

The choice of prior distribution is important in a model comparison since it can potentially have a large impact on the resulting odds-ratio. We want to use astrophysically informed priors wherever possible, or suitable uninformative (but proper) priors otherwise. However, the switching model presents a particular challenge in this respect, as its switching parameters (cf. Section 5.3.2) are ad-hoc and purely phenomenological, and were initially informed by the same data we are trying to test the models on. It is therefore important to avoid potential circularity in properly assessing the prior volume of its parameter space, which affects the relevant ‘Occam factor’ for this model (e.g. see MacKay [116]).

To resolve this, we will make use of the availability of two different and independent data sets: the spin-down and the beam-width data. First, we will perform parameter

estimation using the spin-down data with astrophysical priors where possible and uniform priors based on crude estimates from the data otherwise. For the model parameters common to both the spin-down and beam-width models, we will use the posterior distributions from the spin-down data as prior distributions for the beam-width model. For the remaining beam-width parameters which are not common to both the spin-down and beam-width models we will use astrophysically-motivated priors. In this way we can do model comparisons based on the beam-width data using proper, physically motivated priors. In addition, this enforces consistency between the beam-width and spin-down solutions: for example constraining the two to be in phase.

An obvious alternative is to do the reverse and use the beam-width data to determine priors for the spin-down data. However, for both models, we found difficulties in obtaining good quality posteriors when conditioning on the beam-width data with uniform priors based on crude estimates. Specifically, we found the posteriors to be non-Gaussian and multimodal. To deal with this we would need to use a more sophisticated methodology than that discussed in Appendix 5.A. By contrast this is not the case when conditioning on the spin-down data first (results presented in Section 5.3). This is expected since, even by eye, we see that the spin-down data contains an easily visible ‘signal’, while the beam-width data is relatively ‘noisy’. For this work we are primarily interested in laying out the framework to perform model comparisons and either method should suffice and give the same solution. For now then, we will use the more straight-forward method of using the spin-down data to set priors for the beam-width.

5.3 Defining and fitting the models

In this section we will take each conceptual idea (precession or switching) and define a predictive signal model $f(t|\mathcal{M}_j, \boldsymbol{\theta})$. Each concept may motivate multiple signal models: already we have seen the extension to the original Lyne et al. [111] switching model by Perera et al. [134]. In this work we do not aim to exhaust all known models and are well aware that more models exist that have not yet been considered.

For each concept, we will first discuss the theoretical model, then discuss the choice of priors and finally the resulting posterior and posterior-predictive checks. For both these concepts we build models for both the spin-down and beam-width using the former to inform the priors for the latter as described in Section 5.2.3. Model comparisons will be made on the beam-width data only. In addition to these two concepts, we will also consider a noise-only model for the beam-width data.

It is worth stating that by using the signals-in-noise statistical model, we do not make any assumptions on the cause of the noise other than requiring it to be stationary and Gaussian (cf. Jaynes [86]). Given the uncertain physics of neutron stars and the measurement of pulses, it seems likely the noise will contain contributions both from the neutron star itself, and from the measurement process, with the former dominating. We will add a subscript to the noise component $\sigma_{[\nu, W_{10}]}$ to distinguish between the two data sources.

5.3.1 Noise-only model

5.3.1.1 Defining the noise-only beam-width model

Before evaluating the precession and switching hypothesis, let us first consider a noise-only model. This will introduce some generic concepts and provide a benchmark against which to test other models. The noise-only model asserts that the beam-width data (as seen in panel B of Figure 5.1) does not contain any periodic modulation, but is the result of noise about a fixed beam-width: the signal model $f(t) = W_{10}$ is a constant.

We will not consider the spin-down data under such a hypothesis since it is the beam-width data alone that we will use to make model comparisons and it is clear by eye that such a model is incorrect.

5.3.1.2 Fitting the model to the beam-width data

For the noise-only model we have two parameters which require a prior: the constant beam-width W_{10} and the noise $\sigma_{W_{10}}$. For the beam-width we will set a prior using astrophysical data on the period P from the ATNF database (available at www.atnf.csiro.au/people/pulsar/psrcat, for a description see [117]). This value, $P^{\text{ATNF}} = 0.405043321630 \pm 1.2 \times 10^{-11}$ s, provides a strict upper bound on W_{10} , although typically integrated pulse profiles only occupy between 2% and 10% of the period [112]. Therefore, we will use a uniform prior on $[0, 0.1P^{\text{ATNF}}]$ for $\langle W_{10} \rangle$. The choice of 10% adds a degree of ambiguity into the model comparison since varying it will change the odds-ratio; we investigate this in Section 5.4.3.

For the noise parameter $\sigma_{W_{10}}$ we will use a prior $\text{Unif}(0, 5)$ ms based on a crude estimate from the data. We must be careful here as by doing this we are in a sense using the data twice, but this will not introduce bias into the model comparison provided the same prior is applied for all beam-width models.

The MCMC simulations converge quickly to a normal distribution as shown in Figure 5.2. Of note is the mode of $\sigma_{W_{10}} \sim 2$ ms; this is the Gaussian noise required to explain the variations in W_{10} about a fixed mean. For other models, we hope to explain some of the variations with periodic modulation and the rest with Gaussian noise. So for these models we should expect $\sigma_{W_{10}} < 2$ ms.

In Figure 5.3 we plot the *maximum posterior estimate* (MPE) of the signal alongside the data, i.e. the model prediction when the parameters are set equal to the peak values of the posterior probability distributions. This figure demonstrates that, for the noise-only model, the observed W_{10} has a mean value of approximately 8 ms, and all the variations about this mean are due to the noise. In the following section we will develop models where at least some the variation is explained by periodic modulations.

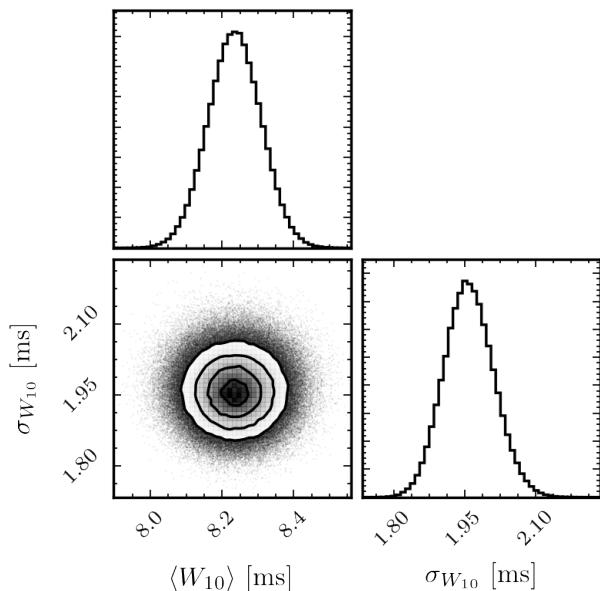


Figure 5.2: The estimated marginal posterior probability distributions for the noise-only model parameters of the beam-width data.

5.3.2 Switching model

The switching idea is phenomenological and we will build the model based on the modification of Lyne et al. [111] by Perera et al. [134]: that is we assume the magnetosphere switches between two meta-stable states *twice* during a single period (the motivation for this was discussed in Section 2.3.3 and we will further motivate it in Section 5.3.2.1). For this work we will assume the switching to be deterministic, although improvements could be made by allowing the switching time to dither, or probabilistic variations in the switching states themselves; see Lyne et al. [111] for some exploration of such ideas. This fully deterministic model captures the primary features without explaining the underlying physics, for example the cause of the switching. Both Jones [89] and Cordes [45] have worked to improve the physical motivations for the switching and provide a consistent picture. Nevertheless, in this work we choose to use the simple phenomenological model as a basis, which can be improved upon in future work.

5.3.2.1 Defining the spin-down rate model

The model proposed by Lyne et al. [111] poses two states for the magnetosphere which we will label as S_1 and S_2 . Then associated to each of these states is a corresponding spin-down rate $\dot{\nu}_1$ and $\dot{\nu}_2$. The smoothly varying spin-down that we observe is a result of the time-averaging process required to measure the spin-down rate. Lyne et al. [111] suggested a square-wave-like switching with a duty rate measuring the fraction of time spent in one state compared to the other. They also proposed a dither in the switching period which will obscure the periodicity, and may give rise to low-frequency structure; we will not consider the dither in this work, but will investigate it in future work. While

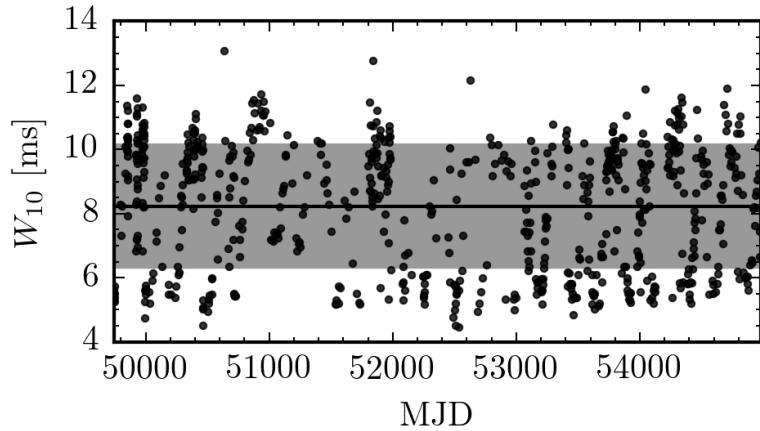


Figure 5.3: Posterior predictive check of the fit of the noise-only model posterior distribution to the data: the solid black line is the maximum posterior estimate (MPE), i.e. the model prediction when the parameters are set equal to the values corresponding to the peaks of the posterior probability distributions. The shaded region indicates the MPE of the $1\sigma_{W_{10}}$ noise about the beam-width model. Black dots are the original data.

studying PSR B0919+06, which also demonstrates a double-peaked spin-down rate like PSR B1828-11, the authors of Perera et al. [134] realised that a (deterministic) switching model which flips once per cycle is incapable of explaining the double-peak observed in the spin-down rate (in particular that one peak is systematically smaller than the other). In order to explain this double-peaked structure, they propose that the mode-changes responsible for switching in the spin-down rate must be doubly-periodic: that is the spin-down rate changes state *twice* during a single cycle. Other modifications, such as introducing a third magnetospheric state, are possible, but in this work we will apply the Perera et al. [134] switching model to PSR B1828-11.

We now discuss the particular formulation of this model used in this analysis, firstly defining the *underlying* spin-down model and then the time-averaging process. To aid in this discussion we plot both the underlying spin-down model and time-average in Figure 5.4 and gradually introduce each feature.

We begin by defining $\tilde{t} = (t - t_{\text{ref}}) + \phi_0 T \bmod (T)$ where $\phi_0 \in [0, 1]$ is an arbitrary phase offset and t_{ref} is a reference time. For all models in this study we will set $t_{\text{ref}} = \text{MJD } 49621$ to coincide with the epoch at which the ATNF database Manchester and Lyne [118] records measurements for PSR B1828-11. Then the function which generates the switching is

$$\dot{\nu}_{\text{u}}(t) = \begin{cases} \dot{\nu}_1 & \text{if } 0 < \tilde{t} < t_1 \text{ or } t_2 < \tilde{t} < t_3 \\ \dot{\nu}_2 & \text{if } t_1 < \tilde{t} < t_2 \text{ or } t_3 < \tilde{t} < T \end{cases}, \quad (5.10)$$

where the subscript ‘u’ denotes that this is the underlying spin-down model.

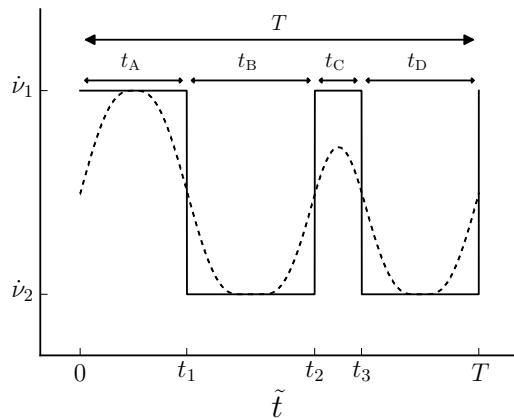


Figure 5.4: Schematic of the doubly-periodic spin-down rate model proposed by Perera et al. [134]. The solid line is the underlying spin-down evolution by the various timescales and initial phase ϕ_0 . This degeneracy may cause difficulties in the MCMC search algorithm, and we therefore fix this gauge freedom by specifying that $|\dot{\nu}_1| < |\dot{\nu}_2|$, $t_A \geq t_C$, and we require that t_A refers to a segment where $\dot{\nu} = \dot{\nu}_1$.

There are multiple ways to parameterise the switching times in the model. For the data analysis, we have chosen to parameterise by the total cycle duration T and three of the segment durations t_A , t_B and t_C .

This model is subject to label-switching degeneracy in the choice of $\dot{\nu}_1$ and $\dot{\nu}_2$ and also between the various timescales and initial phase ϕ_0 . This degeneracy may cause difficulties in the MCMC search algorithm, and we therefore fix this gauge freedom by specifying that $|\dot{\nu}_1| < |\dot{\nu}_2|$, $t_A \geq t_C$, and we require that t_A refers to a segment where $\dot{\nu} = \dot{\nu}_1$.

Based on a cursory inspection of the observed PSR B1828-11 spin-down rate (see Figure 5.1 panel A), it is clear that the secular second-order spin-down rate is non-zero. To model this we will include a constant $\dot{\nu}$ in the underlying spin-down model

$$\dot{\nu}(t) = \dot{\nu}_u(t) + \dot{\nu}(t - t_{\text{ref}}). \quad (5.11)$$

This gives the intrinsic spin-down rate of the pulsar which would be observed if measurements could be taken *without* time-averaging

To simulate the observed spin-down rate, we could time-average Eqn. (5.11) directly. Instead, we choose to mimic the data collection process responsible for the time-averaging. Let us first discuss the data collection process as described by Lyne et al. [111] (this has already been described in other parts of this thesis as the observer-method of calculating the spin-down rate). Observers start with the time-of-arrival of pulsations, which is a measure of the pulsar rotational phase. Taking a 100 day window of data, starting at the earliest observation, a second-order Taylor expansion in the phase is fitted to the data yielding a measurement of $\dot{\nu}$. Then the processes is repeated, sliding the window in intervals of 25 days over the whole data set. The measured $\dot{\nu}$ values at the centre of each window gives a time-averaged spin-down rate.

To mimic this data collection process, we first integrate Eqn. (5.11) twice to generate the phase and then repeat the above process. When integrating, we can ignore the arbitrary phase and frequency offsets since we discard them when calculating the spin-down rate. The resulting spin-down rates constitutes our signal model which is the time-average of Eqn. (5.11). A schematic representation of the sort of spin-down that is then found is given by the dotted curve in Figure 5.4. Clearly, the time-averaged spin-down is much smoother than the underlying spin-down.

It is worth taking a moment to realise that the relation of the time-average spin-down to the underlying model $\dot{\nu}$ depends on both the segment durations and the length of the time average ($t_{\text{ave}} = 100$ days). For the i^{th} segment, if the duration $t_i > t_{\text{ave}}$ then the time-averaged spin-down will ‘saturate’ and have a flat spot as in segment A of the illustration in Figure 5.4. On the other hand, if $t_i < t_{\text{ave}}$ then the maximum spin-down rate in this segment will be a weighted sum of the two underlying spin-down rates as in segment C of the illustration. The weighting is determined by the amount of time the underlying spin-down rate spends in each state during the time-average window.

5.3.2.2 Parameter estimation for the spin-down

In Table 5.1 we list the selected priors. For $\ddot{\nu}$ we define $\ddot{\nu}^{\text{ATNF}} = (8.75 \pm 0.09) \times 10^{-25} \text{ s}^{-3}$ (the value from the ATNF catalogue) and use a normal prior with this mean and standard deviation. In the tables we show the difference with respect to this value. For the remaining parameters we select uniform priors using crude estimates of the data in panel A of Figure 5.1. As previously mentioned, this means we are using the data twice: once in setting up the priors and once for the fitting. Therefore, it would be inappropriate to use the results in a model comparison and this is not our intention: we want to use the posterior distribution as a prior for the beam-width parameter estimation.

Parameter	Distribution	Units
$\dot{\nu}_1$	Unif($-3.66 \times 10^{-13}, -3.64 \times 10^{-13}$)	s^{-2}
$\dot{\nu}_2$	Unif($-3.67 \times 10^{-13}, -3.66 \times 10^{-13}$)	s^{-2}
$\ddot{\nu} - \ddot{\nu}^{\text{ATNF}}$	$\mathcal{N}(0, 9.0 \times 10^{-27})$	s^{-3}
T	Unif(450, 550)	days
t_A	Unif(0, 250)	days
t_B	Unif(0, 250)	days
t_C	Unif(0, 250)	days
ϕ_0	Unif(0, 1)	
$\sigma_{\dot{\nu}}$	Unif(0, 1×10^{-15})	s^{-2}

Table 5.1: Prior distributions for the spin-down switching model.

For the spin-down data under the switching model, the MCMC simulations converge quickly to unimodal and approximately Gaussian distributions. The distributions are plotted in Figure 5.5A, and we summarise the results by their mean and standard-deviation in Table 5.2. In the case of parameters with uniform priors, this indicates that the data is informative and a well defined ‘best-fit’ has been selected. For $\ddot{\nu}$ the

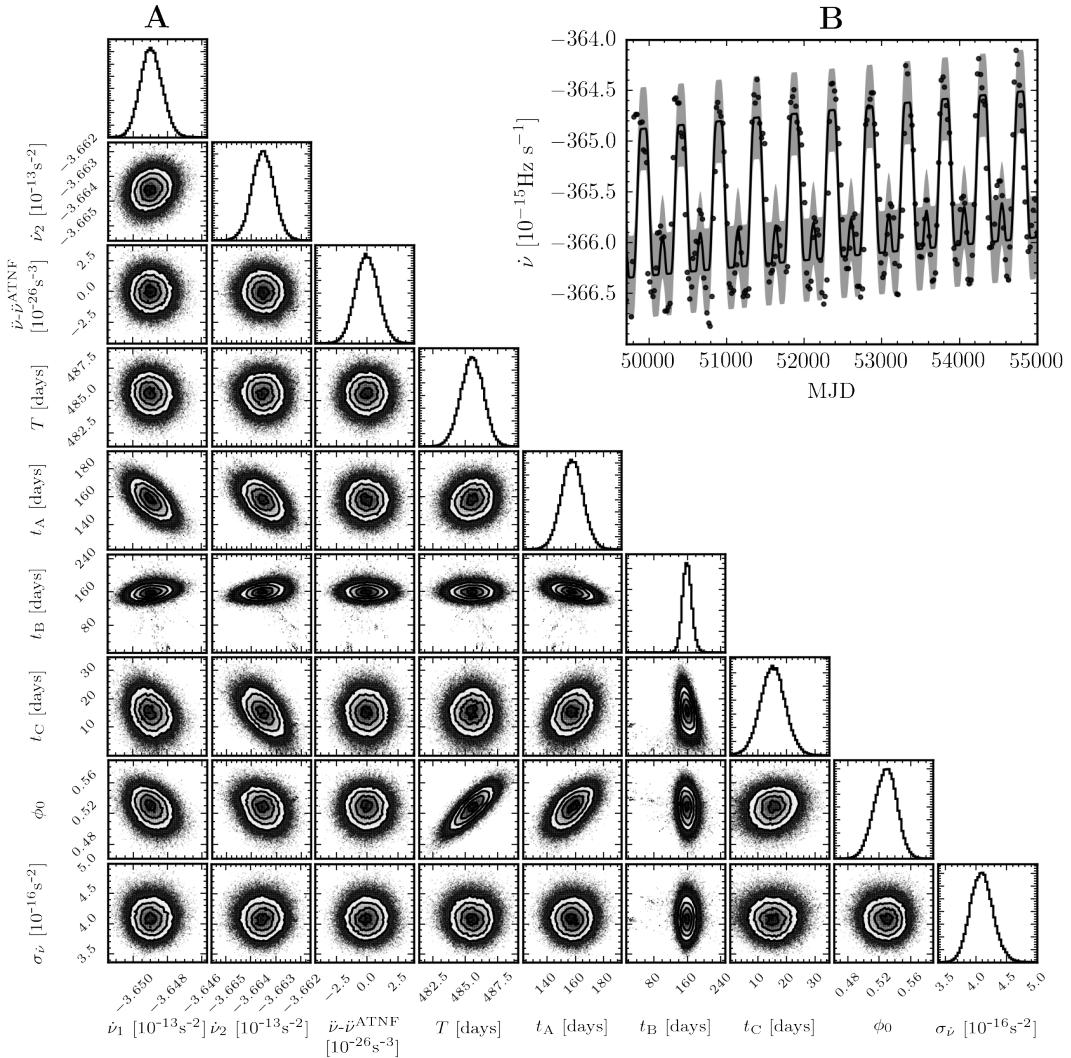


Figure 5.5: **A:** The estimated marginal posterior probability distribution for the Switching spin-down model parameters. **B:** Checking the fit of the model using the maximum posterior values to the data; see Figure 5.3 for a complete description.

Parameter	Mean \pm s.d.	Units
$\dot{\nu}_1$	$-3.6489 \times 10^{-13} \pm 6.33 \times 10^{-17}$	s^{-2}
$\dot{\nu}_2$	$-3.6635 \times 10^{-13} \pm 4.44 \times 10^{-17}$	s^{-2}
$\dot{\nu} - \dot{\nu}^{\text{ATNF}}$	$-3.1051 \times 10^{-28} \pm 9.0 \times 10^{-27}$	s^{-3}
T	485.52 ± 0.8649	days
t_A	157.75 ± 7.6587	days
t_B	159.71 ± 11.7798	days
t_C	15.1379 ± 4.3925	days
ϕ_0	0.5278 ± 0.0143	
$\sigma_{\dot{\nu}}$	$4.0932 \times 10^{-16} \pm 1.84 \times 10^{-17}$	s^{-2}

Table 5.2: Posterior estimates for the spin-down switching model.

posterior has not departed significantly from the (informative) prior meaning that the data agrees with the prior.

To check that our fit is sensible, we plot the observed spin-down data in Figure 5.5B alongside the maximum posterior estimate for the signal. The relative size of the noise component informs us how well the model fits the data: if $\sigma_{\dot{\nu}}$ is of a similar size to the variations in spin-down rate, then the model does poorly and we require a large noise component. In this case the noise-component is smaller than the variations in spin-down rate and the signal model explains most of the variations in the data.

Comparing the maximum posterior values of the four segment times to the baseline on which we time-average (fixed at 100 days) can give an insight into how the model has best fit the data. If we take the posterior of t_C , we find it has a mean value of ~ 15 days which is significantly shorter than the baseline on which we time-average. For the other three segments, their durations are longer than this baseline. The reason that the fit in Figure 5.5B has one maxima smaller than the other, is because the segment duration for that segment, t_C , is shorter than the time-average baseline. This is expected and was precisely the motivation for using the model proposed by Perera et al. [134]; a switching model split into only two segments could not produce this feature.

5.3.2.3 Defining the beam-width model

For the beam-width, we assume that changes in the spin-down rate directly correlate to changes in this beam-width through changes in the beam geometry. Since we require the switching to be doubly-periodic for the spin-down to make sense, so we must require the beam-width to be doubly periodic. That is we define the beam-width model to be

$$W(t) \begin{cases} W_1 & \text{if } 0 < \tilde{t} < t_1 \text{ or } t_2 < \tilde{t} < t_3 \\ W_2 & \text{if } t_1 < \tilde{t} < t_2 \text{ or } t_3 < \tilde{t} < T \end{cases}. \quad (5.12)$$

Lyne et al. [111] noted that the larger beam-widths tended to correlate with the lower (absolute) spin-down rate ($\dot{\nu}_1$ in our model). We could fix this by requiring that $W_1 > W_2$ (recalling that we set $|\dot{\nu}_1| < |\dot{\nu}_2|$), but instead we will not implement such a constraint and allow the data to decide. As with the spin-down, to break the degeneracy in the

times we will require again that $t_A \geq t_C$. We will not assume any secular changes in the beam-width for simplicity.

5.3.2.4 Parameter estimation for the beam-width

Having obtained a sensible fit to the spin-down data, we use the resulting posteriors (as summarised in Table 5.2) to inform our priors for the beam-width data. We can do this only for those parameters common to both the beam-width and spin-down predictions of the switching model: namely the four timescales, and the phase-offset. We would like to relate the spin-down rates $\dot{\nu}_1$ and $\dot{\nu}_2$ to the beam-widths. However, the underlying physics is not understood, and so instead we will take a naive approach and set a prior on the beam-widths from astrophysical data.

For the beam-widths, W_1 and W_2 , we will use the same prior as defined in Section 5.3.1.2 for the noise-only model: namely a uniform prior on $[0, 0.1P^{\text{ATNF}}]$ covering 10% of the spin-period. Using such a prior introduces some ambiguity into the model comparison as the result could be ‘tuned’ by varying the fraction f of the spin-period used to set the uniform prior limits (here $f = 0.1$). This issue is addressed in Section 5.4.3 where we find all sensible choices of f lead to the same overall conclusion.

The final parameter which requires a prior distribution is the noise-component: as described in Section 5.3.1 we apply a prior to $\sigma_{W_{10}}$ using a crude estimate from the data; this is tabulated along with the other priors in Table 5.3.

Parameter	Distribution	Units
W_1	Unif(0, 40.5000)	ms
W_2	Unif(0, 40.5000)	ms
T^*	$\mathcal{N}(485.5, 0.8649)$	days
t_A^*	$\mathcal{N}(158.0, 7.6587)$	days
t_B^*	$\mathcal{N}(160.0, 11.7798)$	days
t_C^*	$\mathcal{N}(15.1379, 4.3925)$	days
ϕ_0^*	$\mathcal{N}(0.5278, 0.0143)$	
$\sigma_{W_{10}}$	Unif(0, 5)	ms

Table 5.3: Prior distributions for the beam-width switching model. Parameters for which the prior is taken from spin-down posteriors are labelled by *.

We plot the posterior estimate in Figure 5.6A which demonstrates non-Gaussianity and multimodal features in the segment times and the phase. This indicates the existence of multiple solutions which could explain the data. We note that the noise component $\sigma_{W_{10}}$ has a mode at 1.6 ms which is less than the 2ms required in the noise-only model. This indicates that some of the variability is being explained by the signal model. In Table 5.4 we summarise the posterior. We find that the posterior modes satisfy $W_1 > W_2$: larger beam-widths are associated with the smaller absolute spin-down rates as found by Lyne et al. [111].

Again we check the predictive power of our estimated posterior by plotting the MPE alongside the data in Figure 5.6B. The fit to the data is not as good as the spin-down

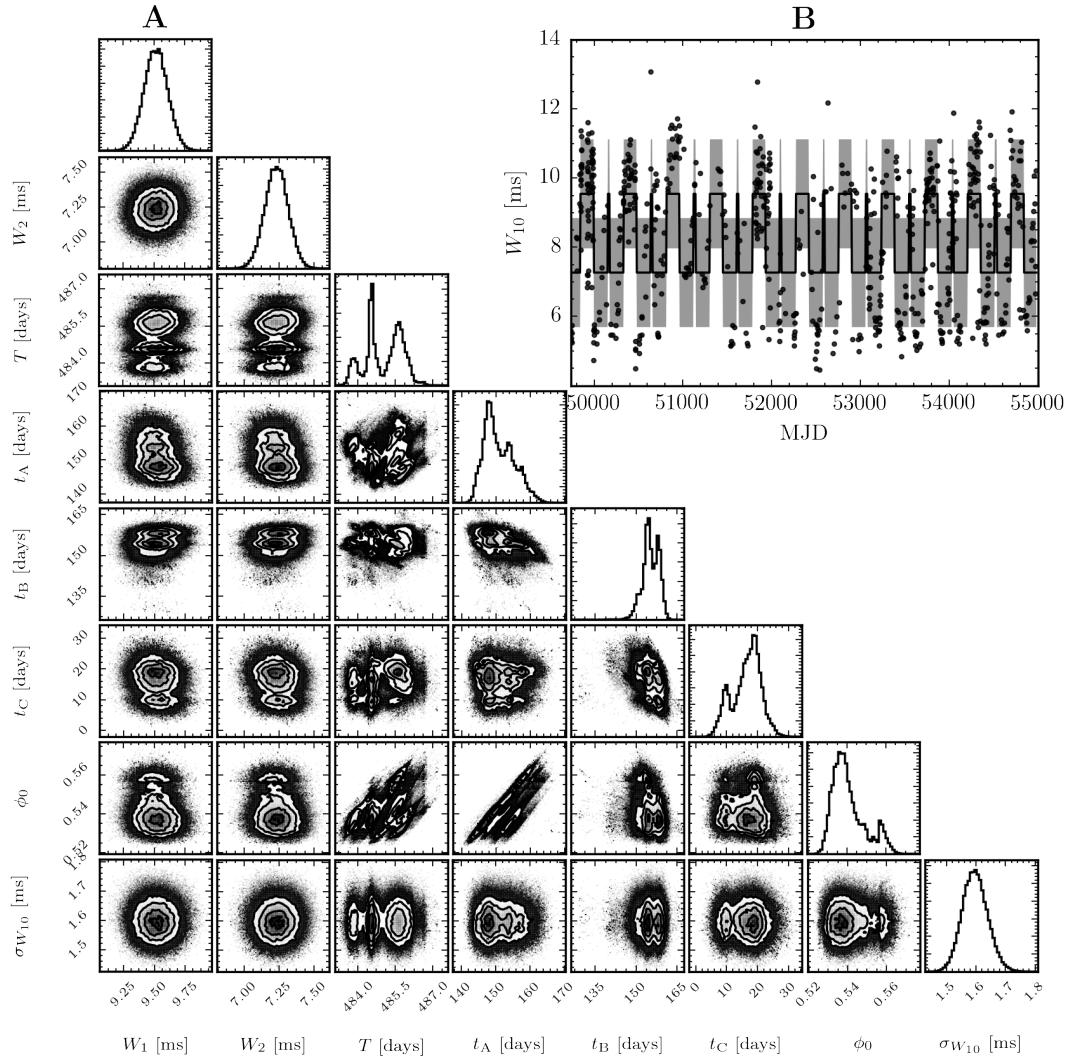


Figure 5.6: **A:** The estimated marginal posterior probability distribution for the Switching beam-width model parameters. **B:** Checking the fit of the model using the maximum posterior values to the data; see Figure 5.3 for a complete description.

Parameter	Mean \pm s.d.	Units
W_1	9.5166 ± 0.0956	ms
W_2	7.2327 ± 0.0830	ms
T	485.04 ± 0.7286	days
t_A	150.98 ± 4.3909	days
t_B	155.29 ± 3.0598	days
t_C	16.4134 ± 4.5771	days
ϕ_0	$0.5409 \pm 8.38 \times 10^{-3}$	
$\sigma_{W_{10}}$	1.5964 ± 0.0427	ms

Table 5.4: Posterior estimates for the beam-width switching model.

fit: by eye it is clear that most data points lie away from the signal model requiring a greater (relative) level of noise.

5.3.3 Precession model

We will now define the precession model and its predictions for the expected signal in the spin-down and beam-width data.

Classical free precession refers to the rotation of a rigid non-spherical body when there is a misalignment between its spin axis and angular momentum vector. For this work we will consider a biaxial star, acted upon by an electromagnetic torque as discussed in the next section. We will work with the angles defined in Jones and Andersson [90]: that is the star emits its EM radiation beam along the magnetic dipole \mathbf{m} which makes an angle χ with the symmetry axis of the moment of inertia, and θ is the so-called wobble angle made between the symmetry axis and the angular momentum vector. We will consider the small-wobble angle regime where $\theta \ll 1$ since this is thought to be the most physical solution for PSR B1828-11. Finally, we define P as the rotation period, and \dot{P} its time derivative, where the small variations due to precession have been averaged over, and

$$\tau_{\text{age}} \equiv \frac{P}{\dot{P}}. \quad (5.13)$$

as a characteristic spin-down age.

5.3.3.1 Defining the spin-down rate model

Observers infer the spin-down rate by measuring the arrival times of pulsations. For a freely precessing star the spin-down rate is periodically modulated on a time-scale known as the *free precession period*, which we will denote as τ_P . This result is referred to as the *geometric modulation* [90] since it is a geometric effect. Under the action of a torque the geometric effect persists, but an additional electromagnetic effect enters owing to torque variations [44]. The authors of Jones and Andersson [90] and Link and Epstein [108] studied both effects in the presence of a vacuum dipole torque [52] and agreed that the electromagnetic contributions dominate for PSR B1828-11: we will therefore neglect the geometric effect. The precession model for PSR B1828-11 has been

developed by Akgün et al. [14] where a non-vacuum dipole torque was considered and additionally by Arzamasskiy et al. [22] where the effect of a plasma-filled magnetosphere was investigated. All of these are potential areas of improvement, but in this work we will restrict our focus to the simplest specification capable of explaining the observations.

For this model comparison, we will use the spin-down rate under an EM torque derived in Section 4.5.1 and verified numerically in Section 4.5.2. This derivation uses a generalisation of the vacuum dipole torque to allow for a braking index $n \neq 3$, but retains the angular dependence; this ansatz may be written as

$$\dot{\nu} = -k\nu^n \sin^2 \Theta, \quad (5.14)$$

where k is a positive constant, and Θ is the polar angle between the dipole and the angular momentum vector as calculated in Eqn. (52) of Jones and Andersson [90]. Then we calculate *secular* solutions in which Θ takes its fixed, time-averaged value. We denote by θ the angle between the symmetry axis of the moment of inertia and the angular momentum vector, and denote by χ the angle between the symmetry axis and the magnetic dipole m . We can then combine the secular solution with an expansion of $\sin^2 \Theta$ in the small θ limit, to give a spin-down rate

$$\dot{\nu}(t) = \frac{1}{\tau_{\text{age}} P} \left(-1 + n \frac{1}{\tau_{\text{age}}} (t - t_{\text{ref}}) + \theta \left[2 \cot \chi \sin(\psi(t)) - \frac{\theta}{2} \cos(2\psi(t)) \right] \right), \quad (5.15)$$

where

$$\psi(t) = 2\pi \frac{t - t_{\text{ref}}}{\tau_P} + \psi_0. \quad (5.16)$$

This is Eqn. (4.68) derived in Section 4.5.1 with the transformation $t \rightarrow t - t_{\text{ref}}$ where, as in the switching model, $t_{\text{ref}} = \text{MJD } 49621$. The first two terms are the secular spin-down rate and its first derivative. The term in the square brackets is the modulation and can be found from appropriate manipulation of Eqn. (58) and (73) in Jones and Andersson [90], or Eqn. (20) in Link and Epstein [108], aside from a factor of χ in the harmonic term which we believe to be a misprint.

For $\chi < \pi/2$ the spin-down rate modulations are sinusoidal. When $\chi \approx \pi/2$ (such that the star is nearly an orthogonal rotator), we will see a strong harmonic at twice the precession frequency. It is precisely this behaviour which is able to explain the doubly-peaked spin-down rate for PSR B1828-11.

5.3.3.2 Parameter estimation for the spin-down

In Table 5.5 we list the priors selected for the spin-down precession model. For τ_P and $\sigma_{\dot{\nu}}$ we use a prior based on a crude estimate from the data. For the spin-down age, braking-index, and pulse period we use a normal prior taking the mean and standard deviation from PSR B1828-11 measurements reported in the ATNF catalogue: the values are listed in Table 5.6. In the analysis, we present the posteriors of the difference to the means of these values for convenience. Finally for ψ_0 we give the full domain of possible values. Since our derivation of the signal models in Section 5.3.3 assumed the

Parameter	Distribution	Units
$\tau_{\text{age}} - \tau_{\text{age}}^{\text{ATNF}}$	$\mathcal{N}(0, 0.3169)$	yrs
$n - n^{\text{ATNF}}$	$\mathcal{N}(0, 0.1700)$	
$P - P^{\text{ATNF}}$	$\mathcal{N}(0, 1.2 \times 10^{-11})$	s
τ_P	Unif(450, 550)	days
θ	Unif(0, 0.1)	rad
χ	Unif($2\pi/5$, $\pi/2$)	rad
ψ_0	Unif(0, 2π)	rad
σ_ν	Unif(0, 1×10^{-15})	s^{-2}

Table 5.5: Prior distributions for the spin-down precession model.

P^{ATNF}	$0.405043321630 \pm 1.2 \times 10^{-11}$ s
n^{ATNF}	16.08 ± 0.17
$\tau_{\text{age}}^{\text{ATNF}}$	213827.91 ± 0.32 yrs

Table 5.6: Measured and inferred values of the precession spin-down model parameters from the ATNF pulsar catalogue [117]. These are given at epoch MJD 49621.

small wobble-angle regime $\theta \ll \chi$ and $\chi \sim \pi/2$, we similarly restrict their uniform priors to the relevant range.

Running the MCMC simulations we plot the resulting posterior in Figure 5.7 and provide a summary in Table 5.7. For all parameters the posterior distribution is Gaussian: in the case of parameters which we gave a uniform prior, this indicates that the spin-down data is informative. For the parameters using an informative prior (from the ATNF catalogue), the posterior and prior are similar. This indicates the data agrees with the prior, but does not significantly improve our estimates.

Parameter	Mean \pm s.d.	Units
$\tau_{\text{age}} - \tau_{\text{age}}^{\text{ATNF}}$	$7.461 \times 10^{-3} \pm 0.3159$	yrs
$n - n^{\text{ATNF}}$	0.0199 ± 0.1701	
$P - P^{\text{ATNF}}$	$1.0436 \times 10^{-14} \pm 1.19 \times 10^{-11}$	s
τ_P	485.56 ± 0.8188	days
θ	0.0490 ± 0.0020	rad
χ	1.5517 ± 0.0013	rad
ψ_0	3.8709 ± 0.0697	rad
σ_ν	$4.0423 \times 10^{-16} \pm 1.81 \times 10^{-17}$	s^{-2}

Table 5.7: Posterior estimates for the spin-down precession model. For the secular spin-down quantities, we report the posterior difference with respect to the values as listed in Table 5.6.

In Figure 5.7B we check the fit of the posterior for the spin-down data. The spin-down model fits to the data points well with only a small amount of noise required to explain the data.

The posterior distributions conditioned on the spin-down data (as summarised in Table 5.7) can be compared with the values reported in Table 2 of Link and Epstein [108]. When comparing, it should be noted that we are considering a longer stretch of data

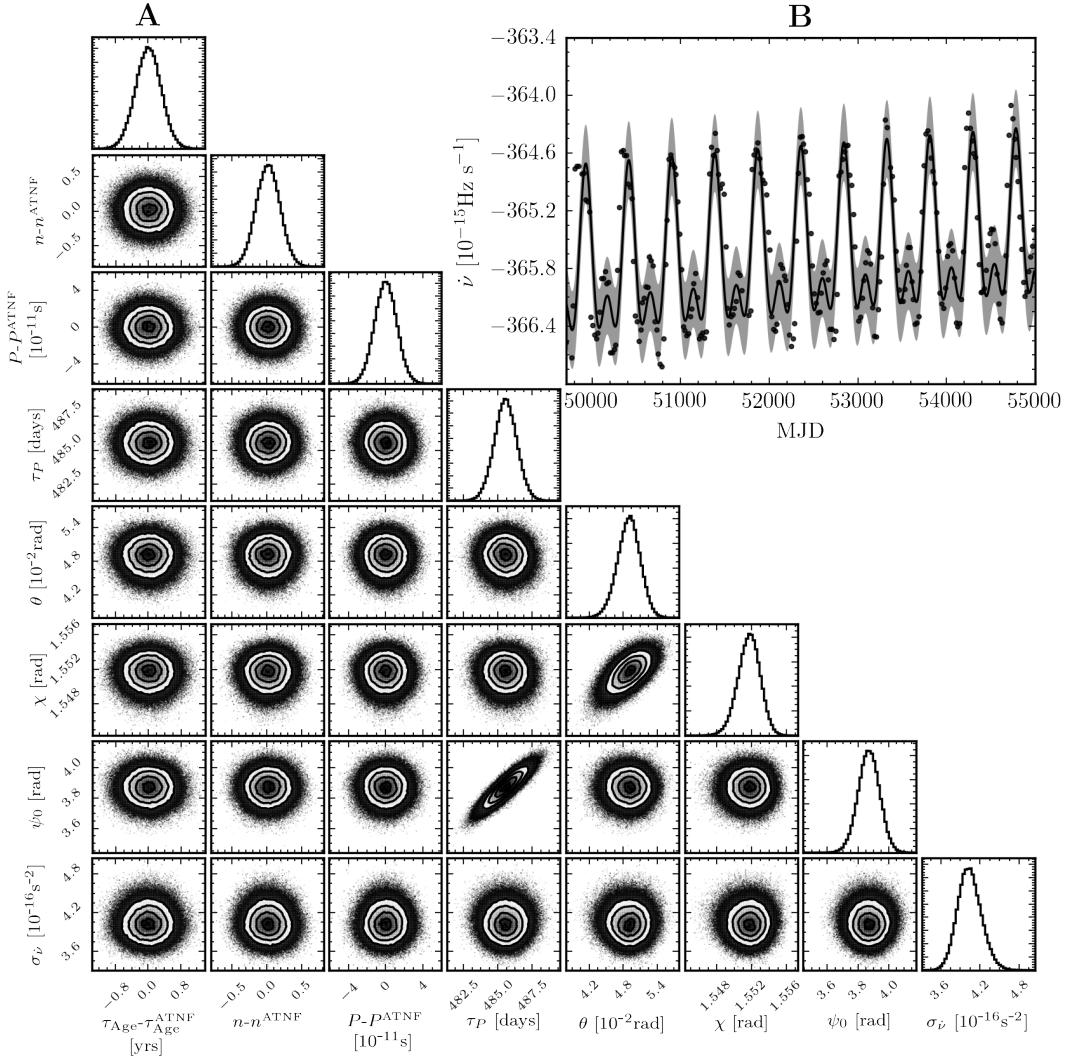


Figure 5.7: **A:** The estimated marginal posterior probability distribution for the precession spin-down model parameters. For the secular spin-down quantities, we show the difference with respect to the values as listed in Table 5.6. **B:** Checking the fit of the model using the maximum posterior values to the data; see Figure 5.3 for a complete description.

which includes most, but not all of the period studied by Link and Epstein [108]. For the two angles χ and θ , the fractional difference is 0.001 and 0.14 respectively while the precession periods differ by a fractional amount 0.05. Clearly the solution found here is similar to that found by Link and Epstein [108].

5.3.3.3 Defining the beam-width model: Gaussian intensity

Modulation of the observed beam due to precession is a purely geometric effect. We have already investigated the beam-width for a Gaussian intensity using our numerical model in Section 4.6.2; in this subsection we will develop the corresponding analytic model of the beam-width under a Gaussian intensity. This is an original development of the precession model, although Akgün et al. [14] have considered the equivalent calculation for the so-called shape parameter S . Ultimately, in agreement with Akgün et al. [14], we will find that a Gaussian intensity profile will not produce the required variations in W_{10} . Instead, in Section 5.3.3.5 we will develop a modified-Gaussian intensity which can explain the variations.

Fixing the beam-axis to coincide with the magnetic dipole \mathbf{m} and following Jones and Andersson [90], we define Θ and Φ as the polar and azimuthal angles of \mathbf{m} with respect to a fixed Cartesian coordinate system with z along the angular momentum vector \mathbf{J} . The observer is fixed in the Cartesian coordinate system with a polar angle ι to \mathbf{J} , and azimuth Φ_{obs} . The slow precessional motion of the spin-vector causes modulation in the angle Θ :

$$\Theta(t) = \cos^{-1} (\sin \theta \sin \chi \sin \psi(t) + \cos \theta \cos \chi), \quad (5.17)$$

which, in the $\theta \ll 1$ limit is approximately

$$\Theta(t) \approx \chi - \theta \sin \psi(t). \quad (5.18)$$

Taking the plane containing the angular momentum vector and the observer, in Figure 5.8 we demonstrate the range of motion of \mathbf{m} over a precessional cycle by a grey shaded region. The region has a mean polar value of χ and a range of 2θ .

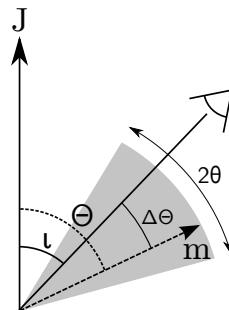


Figure 5.8: Illustration of the angles as the beam-axis \mathbf{m} cuts the plane containing the observer and the angular momentum \mathbf{J} . The grey shaded region indicates the extent to which \mathbf{m} varies over a precessional cycle.

For an observer fixed at an angle ι to \mathbf{J} we define an impact parameter:

$$\Delta\Theta(t) = \Theta(t) - \iota, \quad (5.19)$$

which will vary in time with the precession period τ_P . This impact parameter determines how the observer's line-of-sight cuts the emission beam; if $\Theta(t)$ varies due to changes in $\Delta\Theta(t)$, then the observer will measure the beam to vary on the slow precession time-scale. To help visualise the setup, in Figure 5.9 we plot the unit sphere with points corresponding to the beam-axis \mathbf{m} and the observer. For each of these we have added lines of latitude and longitude. Then we see that $\Delta\Theta$ is the difference between the lines of latitude and we can also define $\Delta\Phi(t) = \Phi(t) - \Phi_{\text{obs}}$ as the difference in the lines of longitude.

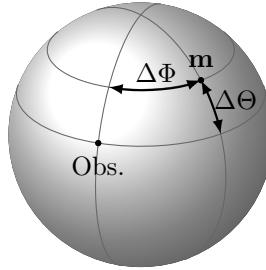


Figure 5.9: The angular position of the observer and the magnetic dipole \mathbf{m} on the unit sphere centred on the star; $\Delta\Theta$ and $\Delta\Phi$ are then the polar and azimuthal angles between them.

The analysis by Stairs et al. [156] characterised the beam by a shape parameter. Link and Epstein [108] used an expansion in $\Delta\Theta$ to model the beam-width and hence the shape-parameter. This allowed them to use their fit to the timing-data to infer the beam-geometry which they found to be hour-glass shaped; see their Figure 5 for a schematic illustration. The authors of Lyne et al. [111] did not use a shape-parameter as it requires time-averaging over a longer base-line, something they wish to avoid in order to be able to observe the switching. Instead, they considered the beam-width at 10% of the maximum, W_{10} , which is measured on a shorter time-baseline (~ 1 hr). If we want to use the beam-width to make a model comparison, we will require a model for W_{10} that is not informed by the data.

The integrated pulse profile of PSR B1828-11 (Figure 4 of Lyne et al. [111]) shows a single peak, often described as core emission [112]. Since we do not have a detailed model of the emission mechanism, we will now consider the most rudimentary and natural beam geometry which fits this: a circularly symmetric (about the beam-axis) intensity which falls-off with a Gaussian function. In the following, we will repeat some of the work of Sec. 4.6.2 for completeness. Specifically, let us define Δd as the central angle between the observer's line-of-sight and the beam (this is the spherical distance between the two points marked in Figure 5.9), then the intensity is

$$\mathcal{I}(t) = \mathcal{I}_0 \exp\left(-\frac{\Delta d(t)^2}{2\rho^2}\right). \quad (5.20)$$

Here \mathcal{I}_0 is the intensity when observed directly along the dipole and ρ measures the angular width of the beam. From the spherical law of cosines, for an observer located at $(\Phi_{\text{obs}}, \iota)$, we have

$$\Delta d(t) = \cos^{-1} [\cos \Theta(t) \cos \iota + \sin \Theta(t) \sin \iota \cos |\Delta \Phi(t)|]. \quad (5.21)$$

The observer will see a maximum pulse intensity at $\Delta \Phi = 0$, given by

$$\mathcal{I}_{\max} = \mathcal{I}_0 \exp \left(-\frac{(\Theta(t) - \iota)^2}{2\rho^2} \right). \quad (5.22)$$

Now let us recognise that Θ varies on the slow precession time-scale, while Φ varies on the rapid spin time-scale: a single pulse consists of Φ varying between $\Phi_{\text{obs}} - \pi$ and $\Phi_{\text{obs}} + \pi$. So over a single pulse, we can treat Θ as a constant. The pulse width W_{10} , as measured by observers, is the duration for which the pulse intensity is greater than 10% of the peak observed intensity. For a single pulse, we can define this duration as the period for which the inequality

$$\mathcal{I} > \mathcal{I}_{\max} \frac{1}{10}, \quad (5.23)$$

is satisfied. Substituting equations (5.20) and (5.22) into (5.23) and rearranging we find that

$$\cos(|\Delta \Phi(t)|) > \frac{\cos \Psi(t) - \cos \Theta(t) \cos \iota}{\sin \Theta(t) \sin \iota}, \quad (5.24)$$

where

$$\Psi(t) = \sqrt{(\Theta(t) - \iota)^2 + 2\rho^2 \ln(10)}. \quad (5.25)$$

Since we treat Θ as a constant over a single pulsation, we can also treat the whole right-hand-side of the inequality as a constant during each pulse.

Now consider a single rotation with the magnetic dipole starting and ending in the antipodal point to the observer's position such that $\Delta \Phi(t)$ increases between $-\pi$ and π during this rotation. Then inequality (5.24) measures the fraction of the pulse corresponding to the beam-width measurement. In terms of the rotation, we define $\delta \Phi$ as the angular width for which the inequality is satisfied and calculate it to be

$$\delta \Phi(t) = 2 \cos^{-1} \left(\frac{\cos \Psi(t) - \cos \Theta \cos \iota}{\sin \Theta \sin \iota} \right). \quad (5.26)$$

Then the beam-width is

$$W_{10}(t) = P \frac{\delta \Phi(t)}{2\pi}, \quad (5.27)$$

from which we arrive at

$$W_{10}(t) = \frac{P}{\pi} \cos^{-1} \left(\frac{\cos \Psi(t) - \cos \Theta(t) \cos \iota}{\sin \Theta(t) \sin \iota} \right). \quad (5.28)$$

In order for the observer to measure the width at 10% of the maximum, the beam intensity must of course drop below this value before increasing again. In reality, we typically observe pulse durations lasting for small fractions of the period, especially when they are close to orthogonal rotators [112].

To set a prior on ρ we consider a special case in which the polar angle of the beam and the observer are at the equator ($\Theta = \iota = \pi/2$). From our spin-down analysis, we know the first of these conditions is true for PSR B1828-11 since χ is close to $\pi/2$. The second condition is based on the assumption that the observer would not see a tightly pulsed beam if they are not close to the polar angle of the beam. In this special instance, inserting Eqn. (5.25) into Eqn. (5.28), the beam-width is

$$W_{10} \Big|_{\Theta=\iota=\pi/2} = \frac{P}{\pi} \sqrt{2 \ln 10} \rho. \quad (5.29)$$

To set a prior on ρ , we can equate this with the beam-widths used in the switching model, for which we set a uniform prior from 0 to $0.1P^{\text{ATNF}}$. To make an even-handed comparison we will therefore set a uniform prior on ρ from 0 to

$$\frac{\pi}{10\sqrt{2 \ln 10}} \approx 0.15, \quad (5.30)$$

so that, for this special case, the prior range of ρ corresponds exactly to the prior range of the beam-widths in the switching model. This prior range will change, but not by orders of magnitude when considering a system close to, but not exactly at, this special case. Therefore, this prior assures that the model comparison does not introduce any significant bias into the model comparison.

5.3.3.4 Parameter estimation for the Gaussian beam-width model

We are in a position to fit the Gaussian beam model to the observed W_{10} values. In Table 5.8 we list the priors taken from the spin-down fit along with three additional priors. For ι we choose a uniform prior in $\cos \iota$ on $[-1, 1]$, this corresponds to allowing ι to range from $[0, \pi]$ (the observer could be in either hemisphere); for ρ we apply the prior from Eqn. (5.30); and for $\sigma_{W_{10}}$ we use a crude estimate based on the data (again we use the same prior for all three models).

Fitting Eqn. (5.28) to the data we discover that the Gaussian beam model is a poor fit to the data. In Figure 5.10B the MPE shows that while the model is able to fit the averaged beam-width, it cannot simultaneously fit the amplitude of periodic modulations.

The posterior distribution (as seen in Figure 5.10A) is Gaussian for all of the parameters except $\cos \iota$ for which it concentrates the probability at $\iota \approx 0$: the observer looks almost down the angular momentum vector. Since $\chi \approx \pi/2$ and $\theta \ll 1$, for each pulsation the beam must therefore sweep out a cone with such a large opening angle it is close to a plane orthogonal to the rotation vector. Meanwhile, the rotation vector is nearly parallel to the angular momentum, since $\theta \ll 1$. As a result, the beam remains approximately orthogonal to the observer for the entirety of each pulsation. We find this result difficult

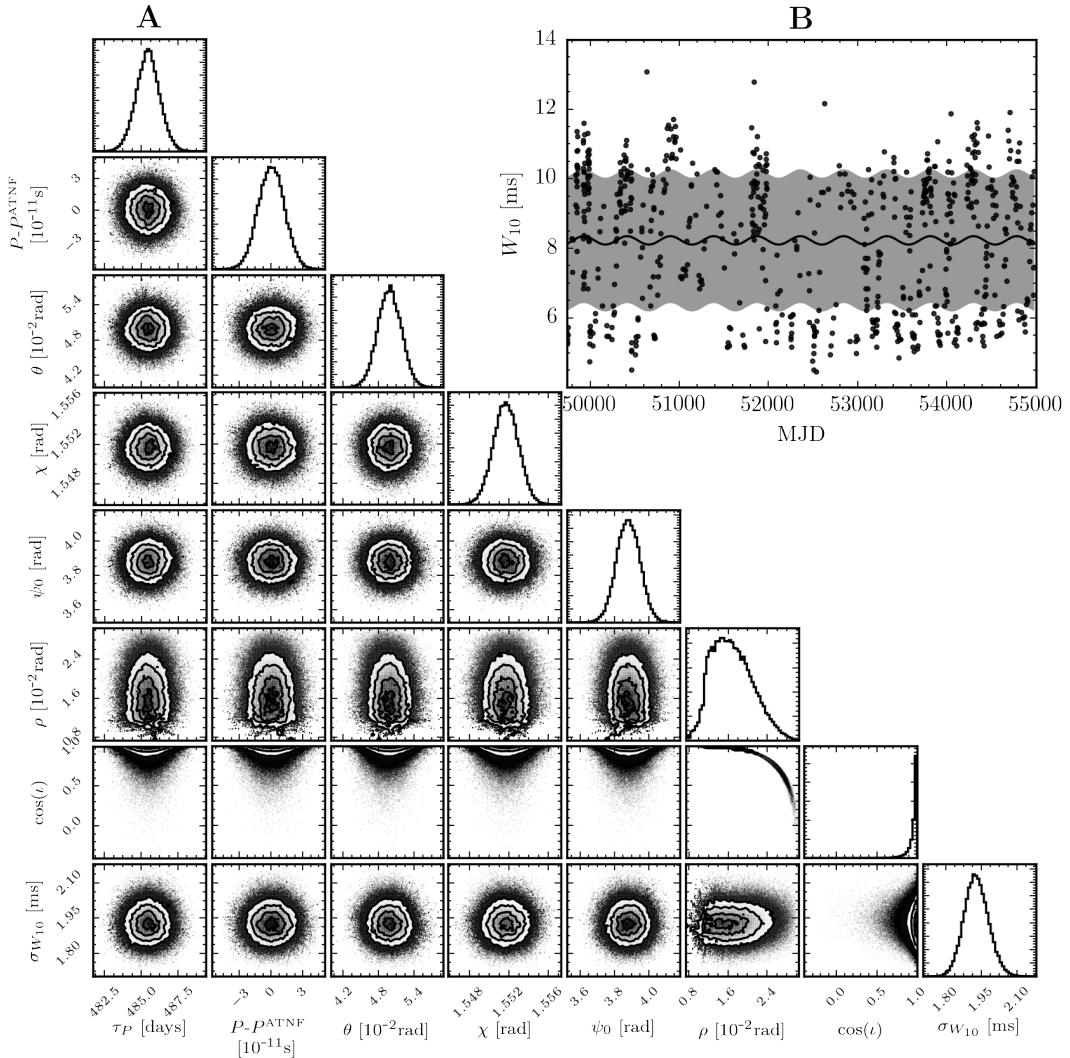


Figure 5.10: **A:** The estimated marginal posterior probability distribution for the Gaussian spin-down model parameters. **B:** Checking the fit of the model using the maximum posterior values to the data; see Figure 5.3 for a complete description.

Parameter	Distribution	Units
τ_P^*	$\mathcal{N}(485.6, 0.8188)$	days
$P - P^{\text{ATNF}} *$	$\mathcal{N}(1.04 \times 10^{-14}, 1.19 \times 10^{-11})$	s
θ^*	$\mathcal{N}(0.0490, 0.0020)$	rad
χ^*	$\mathcal{N}(1.5517, 0.0013)$	rad
ψ_0^*	$\mathcal{N}(3.8709, 0.0697)$	rad
ρ	Unif(0, 0.1500)	rad
$\cos(\iota)$	Unif(-1, 1)	
$\sigma_{W_{10}}$	Unif(0, 5)	ms

Table 5.8: Prior distributions for the beam-width Gaussian precession model. Parameters for which the prior is taken from spin-down posteriors are labelled by *.

to believe on the grounds that the observer would *not* see tightly collimated pulsed emission. For this reason, we conclude that the Gaussian beam-intensity fails to fit the data because the best-fit is unphysical. In retrospect, this result is not surprising since a Gaussian beam intensity is known to have a beam-width (as measured by W_{10}) which is *independent* of the impact angle as discussed by Akgün et al. [14]. This is a direct result of measuring the beam-width with respect to the observed maximum and the self-similar nature of the Gaussian intensity under changes in the impact parameter.

5.3.3.5 Refining the beam-width model: modified-Gaussian intensity

As we have demonstrated that the Gaussian beam is unable to explain both the observed variations and average beam-width, we must now consider how it could be varied in a natural way which does explain the data. One suggestion from Akgün et al. [14] is to impose a sharper cut-off, or introduce a conal component in addition to the Gaussian core emission. We will follow a slightly different path below, one which represents a less drastic modification of the beam profile.

The beam intensity described by Eqn. (5.20) is circularly symmetric about the beam axis as viewed on the surface of the sphere. In the context of the hollow-beam model [143], Narayan and Vivekanand [127] found that pulsar beams can be elongated with the ratio of major to minor axis being ~ 3 for typical pulsars. PSR B1828-11 does not fit into the hollow-beam model (having only a single core component), but nevertheless if the conal emission can be non-circular a generalisation of our core intensity would be to allow for an elliptical beam.

To consider non-symmetric geometries, let us take the planar limit of Eqn. (5.21) by applying small angle approximations in Δd , $\Delta\Theta$ and $\Delta\Phi$:

$$\Delta d(t)^2 = \Delta\Theta(t)^2 + \sin\Theta(t)\sin\iota\Delta\Phi(t)^2. \quad (5.31)$$

This corresponds to setting the observer close to \mathbf{m} in Figure 5.9.

Obviously $\Delta\Phi$ ranges over $[0, 2\pi]$ in each rotation, but when $\Delta\Phi$ is not small, the intensity vanishes rapidly due to the Gaussian beam shape Eqn. (5.22). Therefore,

Eqn. (5.31) is a good approximation for the separation when the beam is pointing near to the observer, while away from this it is a poor approximation, but the intensity is negligible and so the differences are inconsequential.

We can now allow for an elliptical beam geometry by postulating the beam intensity to be

$$\mathcal{I}(t) = \mathcal{I}_0 \exp \left(-\frac{-\Delta\Theta(t)^2}{2\rho_1^2} - \frac{(\sin\Theta(t)\sin\iota\Delta\Phi(t))^2}{2\rho_2^2} \right). \quad (5.32)$$

Then to calculate the beam-width, we first find the maximum:

$$\mathcal{I}_{\max} = \mathcal{I}_0 \exp \left(-\frac{-\Delta\Theta(t)^2}{2\rho_1^2} \right). \quad (5.33)$$

Solving for the beam-width we find

$$W_{10}(t) = \frac{P}{\pi} \frac{\sqrt{2 \ln 10} \rho_2}{\sin\Theta(t)\sin\iota}, \quad (5.34)$$

which is independent of ρ_1 , the latitudinal standard-deviation. The extra degree of freedom introduced in Eqn. (5.32) is irrelevant to the beam-width measure because W_{10} is defined by the ratio of the intensity to that at the observed peak \mathcal{I}_{\max} .

This loss of a degree of freedom means that Eqn. (5.34) is an equivalent to an expansion of Eqn. (5.28) in the planar limit (i.e. the non-circular nature introduced by Eqn. (5.32) does not manifest in the prediction for W_{10}) and so will suffer the same problems if fitted to the data. To further generalise our intensity model we will therefore modify the beam-geometry by allowing a varying degree of non-circularity. This is done by expanding the longitudinal standard deviation as

$$\rho_2(t) = \rho_2^0 + \rho_2''\Delta\Theta(t)^2. \quad (5.35)$$

Note that we have neglected to include a linear term here, forcing the geometry to be longitudinally symmetric about the beam-axis. Preliminary studies began by fitting a linear term only (this giving a modulation at the frequency $1/\tau_P$), but it was found that including a second-order term (which provides modulation at both $1/\tau_P$ and $2/\tau_P$) gave a better fit. Including both terms, we found that the data was unable to provide inference on both ρ_2' and ρ_2'' due to degeneracy. In light of this, we drop the first term, but keep the second, which we feel is the simplest model which is able to fit the data.

Solving for the beam-width (i.e. with Eqn. (5.35) substituted into Eqn. (5.32)) we obtain a signal model

$$W_{10}(t) = \frac{P}{\pi} \sqrt{\frac{2 \ln 10}{\sin\Theta(t)\sin\iota}} (\rho_2^0 + \rho_2''\Delta\Theta(t)^2), \quad (5.36)$$

which we will refer to as the *modified-Gaussian precession beam-width* model.

5.3.3.6 Parameter estimation for the modified-Gaussian precession beam-width

For Eqn. (5.36), we give the relevant prior distributions in Table 5.9. As in the previous Gaussian model, we let ι range over $[0, \pi]$; for ρ_2^0 , we apply the prior on intensity widths as given by Eqn. (5.30); and for ρ_2'' we will use a normal prior with zero mean favouring a Gaussian intensity. The *standard-deviation* of this prior can have a measurable impact on the inference: if it is too small then the degree of freedom introduced by Eqn. (5.35) is effectively removed. Instead, we want to make it significantly larger than the (a priori unknown) posterior value of ρ_2'' : this generates a so-called non-informative prior. To set the prior standard-deviation then, we need to provide a rough scale for what value ρ_2'' should have. To do this we will define our prior expectation such that

$$\rho_2(\Delta\Theta = \rho_2^0) \sim 2\rho_2^0, \quad (5.37)$$

which is to say we expect ρ_2 to increase by no more than a factor of order unity over angular distances of the beam-width comparable to ρ_2^0 (the beam-width when the observer cuts directly through the beam-axis). This amounts to assuming that the beam does not depart very far from circularity. Plugging this into Eqn. (5.35), we get

$$\rho_2'' \sim \frac{1}{\rho_2^0}. \quad (5.38)$$

From this, we use the upper limit from the uniform prior on ρ_2^0 (as calculated in Eqn. (5.30)), to set the standard-deviation for ρ_2'' at $1/0.15 \approx 7$. We also tested different choices of ρ_2'' and found that the posteriors and odds-ratios were robust to the choice, provided the standard-deviation did not exclude the posterior value reported in Table 5.10.

Parameter	Distribution	Units
τ_P^*	$\mathcal{N}(485.6, 0.8188)$	days
$P - P_{\text{ATNF}}^*$	$\mathcal{N}(1.04 \times 10^{-14}, 1.19 \times 10^{-11})$	s
θ^*	$\mathcal{N}(0.0490, 0.0020)$	rad
χ^*	$\mathcal{N}(1.5517, 0.0013)$	rad
ψ_0^*	$\mathcal{N}(3.8709, 0.0697)$	rad
ρ_2^0	Unif(0, 0.1464)	rad
ρ_2''	$\mathcal{N}(0, 6.8308)$	rad ⁻²
$\cos(\iota)$	Unif(-1, 1)	
$\sigma_{W_{10}}$	Unif(0, 5)	ms

Table 5.9: Prior distributions for the beam-width modified-Gaussian precession model. Parameters for which the prior is taken from spin-down posteriors are labelled by *.

The MCMC simulations converge quickly to a Gaussian distribution as shown in Figure 5.11A and the posterior is summarised in Table 5.10. The model parameters common to the spin-down model do not vary significantly from the spin-down posterior: this indicates the two models are consistent. We find that ι is close to $\pi/2$ as expected, ρ_2 is sufficiently small indicating a narrow pulse beam, but ρ_2'' has departed from its

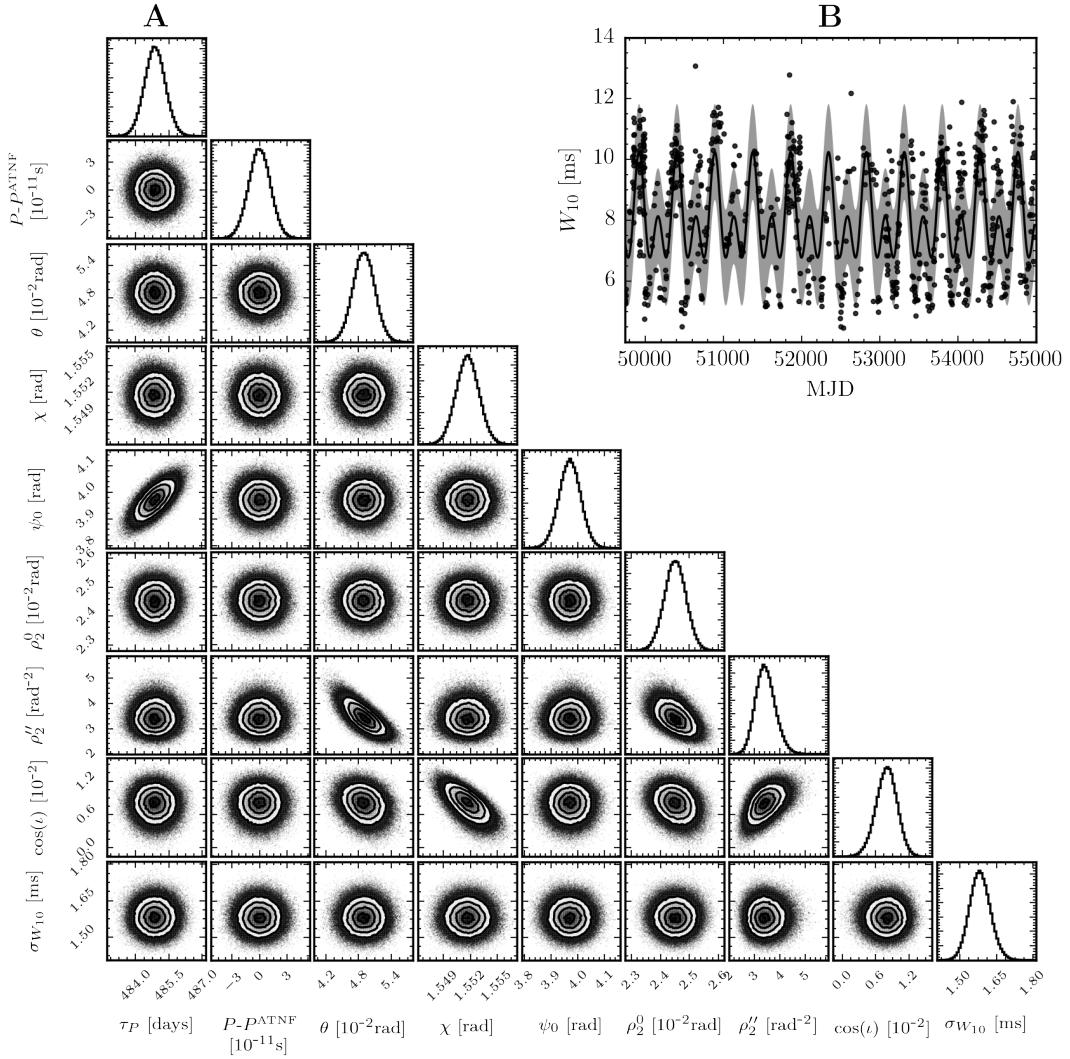


Figure 5.11: **A:** The estimated marginal posterior probability distribution for the modified-Gaussian precession beam-width model parameters. **B:** Checking the fit of the model using the maximum posterior values to the data; see Figure 5.3 for a complete description.

prior mean of zero. This confirms that our generalisation of the Gaussian intensity, Eqn. (5.35), is important in fitting the data.

In Figure 5.11B we perform the posterior predictive check plotting the MPE alongside the data. This demonstrates that the best-fit puts χ within θ of ι such that during the precessional cycle the beam-axis passes twice through observer's location. This corresponds to the grey region in Figure 5.8 intersecting the observer's line-of-sight. When this happens, the modulation of the beam-width picks up a second harmonic at twice the precession frequency. The minima in Figure 5.11B corresponds to the point in the precessional phase when the beam-axis point directly down the observer's line-of-sight.

Parameter	Mean \pm s.d.	Units
τ_P	484.87 ± 0.4706	days
$P - P^{\text{ATNF}}$	$-1.7719 \times 10^{-13} \pm 1.19 \times 10^{-11}$	s
θ	0.0490 ± 0.0020	rad
χ	1.5517 ± 0.0013	rad
ψ_0	3.9701 ± 0.0403	rad
ρ_2^0	0.0245 ± 0.0004	rad
ρ_2''	3.4421 ± 0.3878	rad $^{-2}$
$\cos(\iota)$	$7.9326 \times 10^{-3} \pm 1.9 \times 10^{-3}$	
$\sigma_{W_{10}}$	1.5833 ± 0.0422	ms

Table 5.10: Posterior estimates for the beam-width modified-Gaussian precession model.

5.3.3.7 Recreating the beam-geometry

Since we have defined a beam-intensity in Eqn. (5.32) we can recreate the beam-geometry and pulse-shape from our MPE values. The data we have does not provide information about the latitudinal beam-shape parameter ρ_1 ; therefore we consider that there are a family of beam-geometries parameterised by $\rho_1 = \lambda\rho_2^0$ where λ is an arbitrary scale parameter and ρ_2^0 is the MPE value.

In Figure 5.12 we pick four illustrative values for λ and plot the resulting beam-geometry as contour lines at fixed fractions of the maximum beam intensity (which occurs at the origin). This demonstrates that the beam-geometry has an hour-glass shape in agreement with Link and Epstein [108], although this becomes weaker with smaller values for λ .

In Figure 5.12, a pulse corresponds to a horizontal cut through the intensity at fixed $\Delta\Theta$. Our posterior distribution, Figure 5.11A, also provides information on how the observations cut through this beam-geometry. Under the precession hypothesis, the observer has a time-averaged $\Delta\Theta$ of $\chi - \iota$: this has been plotted as a horizontal dashed line in Figure 5.12. Precession modulates $\Delta\Theta$ about this average value by $\pm\theta$; the observer's line-of-sight through the beam therefore varies by $2\theta \approx 0.1$ rad over a precessional cycle. We have plotted a grey shaded region in Figure 5.12 to show the extent, $\chi - \iota \pm \theta$, over which $\Delta\Theta$ varies during a precessional cycle.

We stress here that the contour lines cannot be used directly to measure the beam-width W_{10} . This is because W_{10} is defined as the width at 10% of the peak intensity for that observed pulse and not the maximum intensity of the beam. The peak intensity for an observed pulse (a horizontal slice) is the intensity at $\Delta\Phi = 0$ and it is with respect to this, which W_{10} is measured.

By construction, the four beam geometries in Figure 5.12 all produce the same W_{10} behaviour as observed in Figure 5.11B. The reason for this is that we have lost information on the total intensity by using W_{10} ; other measurements of the beam-width could potentially yield more information and better constrain the beam geometry.

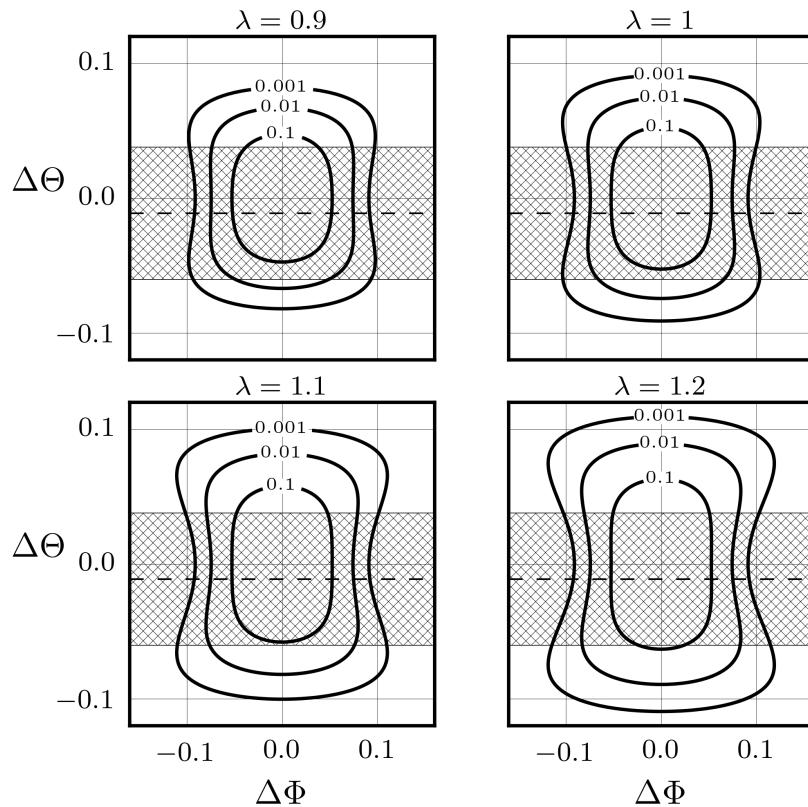


Figure 5.12: Recreating the beam geometry from the MPE of the modified-Gaussian precession beam-width model parameters for four different values of λ . Thick black lines indicate contour lines of the intensity function at fractions of the maximum intensity. The hatched area indicates the region of horizontal cuts (pulses) sampled by the observer: this has a mean, $\chi - \iota$, close to zero (marked by a dashed line) and varies by $\pm\theta$ about this mean.

Fixing $\lambda = 1$ we can also consider the variations in the pulse profile. In Figure 5.13 we plot the normalised intensity for three values of $\Delta\Theta$ corresponding to the mean, and edges of the grey region in Figure 5.12. This figure shows that the narrow beam-widths occur when $\Delta\Theta$ is small, which, since χ is close to $\pi/2$ coincide with the larger (absolute) spin-down rates. This agrees with the findings of Lyne et al. [111] and this figure can be directly compared with panel C in Figure 3 of that work.

5.4 Estimating the odds-ratio

5.4.1 Thermodynamic integration

Having checked that our MCMC simulations are a reasonable approximation to the posterior distribution we now calculate the marginal likelihood for each model and then their odds-ratio. To calculate the marginal likelihood we will use thermodynamic integration. This requires running N parallel MCMC simulations and raising the likelihood to a power $1/T$ where T is the ‘temperature’ of the chain. This method was originally

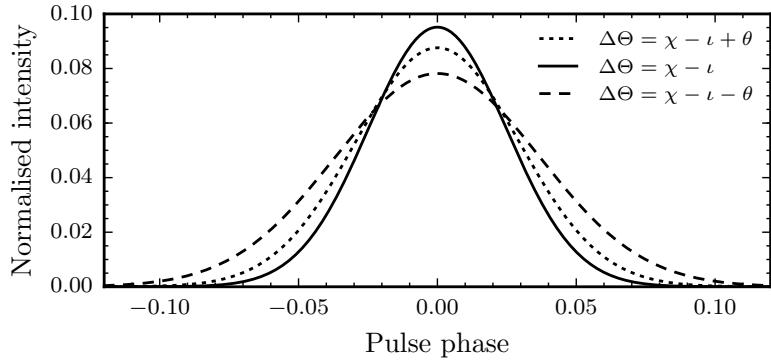


Figure 5.13: Recreating the pulse profiles for three particular slices through the beam using a fixed value of $\lambda = 1$.

proposed by Swendsen and Wang [158] to improve the efficiency of MCMC simulations for multimodal distributions. In this work we use this method not to help with the efficiency of the simulations*, but instead so that we can apply the method prescribed by Goggans and Chi [67] to estimate the evidence as follows.

First we define the inverse temperature $\beta = 1/T$, then let the marginal likelihood as a function of β be

$$Z(\beta) = \int P(\text{data}|\boldsymbol{\theta})^\beta P(\boldsymbol{\theta})d\boldsymbol{\theta}. \quad (5.39)$$

When $\beta = 1$, this gives exactly the marginal likelihood first defined in Eqn. (5.3). After some manipulation we see that

$$\frac{1}{Z} \frac{\partial Z}{\partial \beta} = \frac{\int \ln(P(\text{data}|\boldsymbol{\theta}))P(\text{data}|\boldsymbol{\theta})^\beta P(\boldsymbol{\theta})d\boldsymbol{\theta}}{\int P(\text{data}|\boldsymbol{\theta})^\beta P(\boldsymbol{\theta})d\boldsymbol{\theta}}. \quad (5.40)$$

From this, we note that the right-hand-side is an average of the log-likelihood at β and so

$$\frac{\partial}{\partial \beta} (\ln(Z)) = \langle \ln(P(\text{data}|\boldsymbol{\theta})) \rangle_\beta. \quad (5.41)$$

Using the likelihoods calculated in the MCMC simulations, we numerically integrate the averaged log-likelihood over β which yields an estimation of the marginal likelihood. To be confident that the estimate is correct, we ensure that we use a sufficient number of temperatures and that they cover the region of interest.

5.4.2 Results

Applying the thermodynamic integration technique to all the models, we estimate the evidence for each model. Taking the ratio of the evidences gives us the Bayes factor and since we set the ratio of the prior on the models to unity, the Bayes factor is exactly the odds ratio (see Eqn. (5.2)).

*All the posteriors are either unimodal or multimodal with little separation between the modes.

We present the \log_{10} odds-ratio between the models in Table 5.11. A positive value indicates that the data prefers model A over model B. Note that the error here is an estimate of the systematic error due to the choice of β values [see 62, for details].

Model A	Model B	$\log_{10}(\text{odds-ratio})$
switching	noise-only	57.4 ± 0.5
precession*	noise-only	60.1 ± 0.5
precession*	switching	2.7 ± 0.5

Table 5.11: Tabulated log-odds-ratios for all models. *By the precession model here we mean the precession with a modified Gaussian beam model as discussed in Section 5.3.3.5.

This table allows quantitative discrimination amongst the models. The first two rows compare the switching and modified-Gaussian precession models against the noise-only model with the periodic modulating models being strongly preferred in both cases. Then in the last row we present the log-odds-ratio between the modified-Gaussian precession and switching model which shows that the data prefers the precession Modified Gaussian model by a factor $10^{2.7}$. Using the interpretation of Jeffreys [87], the strength of this evidence can be interpreted as ‘decisive’ in favour of this precession model. For completeness, we also mention that the odds-ratio for the non-modified Gaussian model (which failed to fit the data in a physically meaningful way) against the noise-only model was 3.1 ± 0.6 .

5.4.3 Effect of the choice of prior

For both beam-widths in the switching model we used uniform priors on $[0, fP^{\text{ATNF}}]$ with $f = 0.1$ and these were transformed to also provide a fair prior on ρ_2^0 and ρ_2'' . This choice of f was taken from the upper limit quoted in Lyne and Manchester [112] for typical values of the pulse width. Nevertheless, changing f can have a measurable impact on the odds-ratio and so we will now study this effect.

To begin, we rewrite Eqn. (5.3), the marginal likelihood, by factoring out the N parameters which have a uniform prior

$$P(\text{data}|\mathcal{M}_i) = \frac{1}{\prod_i^N (b_i - a_i)} \int P(\text{data}|\boldsymbol{\theta}, \mathcal{M}_i) P(\boldsymbol{\theta}^*|\mathcal{M}_i) d\boldsymbol{\theta}, \quad (5.42)$$

where by $P(\boldsymbol{\theta}^*|\mathcal{M}_i)$ we mean the probability distribution of all remaining parameters which are not factored out, and $[a_i, b_i]$ is the range for the i^{th} uniform parameter. For the switching beam-width, the prefactor of this integral (factoring out the prior on W_1 and W_2) is $(fP^{\text{ATNF}})^{-2}$: varying f directly impacts the evidence for the switching model. For the precession model, we cannot factor the dependence on f in the same way, as we use a central normal prior on ρ_2'' . We set the standard-deviation of this prior by applying Eqn. (5.38) so that it is inversely proportional to f . If both the prior on ρ_2^0 and ρ_2'' had been proportional to f we would have an exact cancellation in the odds-ratio and hence

no dependence on f . This is not the case and due to our prior on ρ_2'' the odds-ratio will depend on f .

To test the dependence, in Figure 5.14 we plot the log odds-ratio as a function of $\log_{10}(f)$ (note that $f = 0.1$ corresponds to the result in Table 5.11). There are several features to understand. First, for $f \lesssim 0.024$ the odds-ratio rapidly grows, favouring precession; this is because for such small values of f , the beam-width switching prior excludes the values of W_1 required to fit the data. As a result the switching solutions are unnaturally disadvantaged compared to the precession solutions. Such odds-ratios do not fairly compare the models.

For $0.024 \lesssim f \lesssim 0.3$ the log-odds-ratio is approximately linear growing from 1.44 when $f = 0.024$ to 3.21 when $f = 0.3$. In this region the solutions for both models are supported by the prior in that it does not exclude or disfavour the posterior value. The variation in the odds-ratio results from changes in the prior volume of the switching model, the evidence for the precession model is constant in this region. Small f values maximally constrains the prior volume for the switching model (without excluding posterior values) and hence give the greatest weight of evidence to switching and the smallest odds-ratios. For larger f values the log of the prior volume grows linearly resulting in the observed growth.

For $f \gtrsim 0.3$ our choice of standard-deviation for ρ_2'' starts to disfavour the posterior value because it is inversely proportional to f . As a result the evidence for the precession model decreases faster than the loss of evidence for the switching model leading to the observed drop in the odds-ratio. In this case it is the precession solutions which are unnaturally disadvantaged by our choice of prior and so, as in the $f \lesssim 0.024$ case, we do not consider such odds-ratios as a fair comparison of the models.

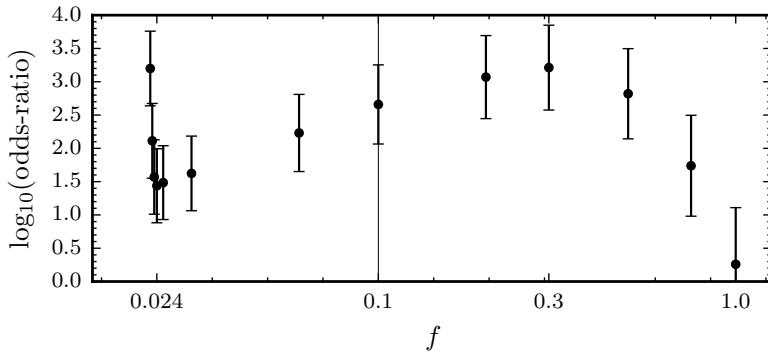


Figure 5.14: Dependence of the odds-ratio with f , the fraction of P^{ATNF} used to constrain the beam-width priors. The vertical line marks the choice of $f = 0.1$ used in our model comparison based on the upper limit given by Lyne and Manchester [112].

In summary the log-odds-ratio and hence our conclusion is robust to reasonable variations in f from 0.05 to 0.5.

5.5 Discussion

In this chapter, we are using a data set (provided by Lyne et al. [111]) on the spin-down and beam-width of PSR B1828-11 to compare models for the observed periodic variations. The two concepts under consideration are free precession and magnetospheric switching. In order to be quantitative, we built signal models for the beam-width and spin-down from these conceptual ideas. Using the spin-down data to create proper, physically motivated priors for the beam-width parameters, we then perform a Bayesian model comparison between the models asking ‘which model does the beam-width data support?’. For the models considered here, the data most strongly supports a precession model with a modified-Gaussian beam geometry allowing for an elliptical beam where the ellipticity has a latitudinal dependence.

To be clear, this does not rule out the switching interpretation since we have not tested an exhaustive set of models — we can only compare between particular models. As an example we could imagine modifying the switching model such that either the switching times, or the magnetospheric states are probabilistic (or a combination of the two). Further we believe there is good grounds to develop models combining the precession and switching interpretation like those discussed in Jones [89] and the preliminary results of Section 4.7.

In addition to the data considered here, a number of high-time-resolution observations of PSR B1828-11 were performed by the Parkes telescope, as discussed in Stairs et al. [157] and Lyne [109]. This data set shows interesting variability in beam-width on short timescales of $\mathcal{O}(100)$ pulses. While the qualitative ‘noisiness’ of the beam-width data is already apparent from the current data-set (e.g. see Figure 5.1), such high-time-resolution data could be very interesting to include in a more detailed future model comparison.

The process of fitting the models to the data and performing posterior predictive checks also provides a mechanism to evaluate the models. For both spin-down models the maximum posterior plots with the data (Figs. 5.5B and 5.7B) revealed a systematic failure to fit the second (slightly lower) minima. This suggests new ingredients could be introduced to both models to explain this.

The posteriors for the precession model indicate that PSR B1828-11 is a near-orthogonal rotator and we observe it from close to the equatorial plane. If this is the case, and the two beams of the pulsar are symmetric about the origin, then we expect to see the second beam as an interpulse. Indeed, we discuss further in Appendix 5.B how during the precessional cycle we should expect the intensity of this second beam to dominate at certain phases. Since no such interpulse is reported, either the second beam is weaker, or the beams must have a kink of greater than 4.6° (see Appendix 5.B).

In this chapter, we have developed the framework to evaluate models for the variations observed in PSR B1828-11. This is not intended as an exhaustive review of all models,

but rather a discussion on the intricacies that arise such as setting up proper and well-motivated priors. This lays the groundwork for a more exhaustive test of all available models and can also be extended by including other data sources.

Appendix 5.A Procedure for MCMC parameter estimation

The procedure used to simulate the posterior distribution can determine the quality of the estimation. Therefore, we will now set out an algorithmic method to ensure that our results are reproducible.

To estimate the posterior given a signal model and prior, we run two MCMC simulations: an initialisation and production. In the following the term *walker* refers to single chains in the MCMC simulation. The Foreman-Mackey et al. [62] implementation runs a number of these in parallel for each simulation.

- For the initialisation run, we draw samples from the prior distribution to set the initial parameters for each walker. The simulation therefore has the chance to explore the entire parameter space. After a sufficient number of steps, the walkers will converge to the local maxima in the log-likelihood. By visually inspecting the data we determine the nature of the local maxima: in all cases a single maxima dominated such that, given a sufficient length of simulation, we expect all walkers to converge to this maxima. Alternatively we could have found multiple similarly strong maxima, in this case further analysis would be required. This was not found to be the case for any of the models in this analysis.
- For the second step we set the initial state of 100 walkers by uniformly dispersing them in a small range about the maximum-likelihood found in the previous step. The simulation proceeds from this initial state and we divide the resulting samples equally into two: discarding the first half as a so-called ‘burn-in’. We retain the second half as the production data used to estimate the posterior. The burn-in removes any memory of the artificial initialisation of the walkers at the start of this step.

Having run an MCMC simulation we check that the chains have properly converged (for a discussion on this see Chapter 10 of Gelman et al. [64]). The MCMC simulations provide an estimate of the posterior densities for the model parameters. We will also perform ‘posterior predictive checks’ to ensure the posterior is a suitable fit to the data, i.e. we compare the data to the model prediction when the model parameters are set to the values corresponding to the peaks of the posterior probability distributions.

Appendix 5.B Implications for the unobserved beam

The precession model developed here assumes the observer only ever sees one pole of the beam-axis, but in the canonical model we often imagine there is also emission from an opposite magnetic pole. In several pulsars this can be seen as an *interpulse*, 180° out of phase from the main pulse [112, 115]; these pulsars are generally found to be close to orthogonal rotators[†]. No such interpulse is reported for PSR B1828-11. We will now discuss the implications of this given the precession interpretation.

Let us imagine a scenario where the observer is in the northern (magnetic) hemisphere and label the beam protruding into their hemisphere (which they will see with the greater intensity due to the smaller angular separation) at the start of the thought experiment as the north pole. Then, when $\Theta < \pi/2$, the north and south pole make angles Θ and $\pi - \Theta$ respectively with the fixed angular momentum vector. Now we see that if, during the course of the precessional cycle, $\Theta > \pi/2$ then the south pole will protrude into the northern hemisphere and the north pole into the southern hemisphere. Provided both poles are identical, but regardless of the details of the beam-geometry, at this time we must expect the observer to see the south pole at greater intensity than the north pole. An example of this is shown in Figure 5.15, but note that the observer will see the greatest intensity from the south-pole half a rotation after this instance.

Our posterior distributions inform us that, if the precession interpretation is correct, we are in exactly this situation: Θ ranges from 85.8° to 92.3° over a precessional cycle[‡] so we should see the interpulse.

This is readily explained if the south pole is substantially weaker in intensity, or by the one-pole interpretation Manchester and Lyne [118]. Alternatively, it could be that the two beams are not diametrically opposed but are latitudinally ‘kinked’. In the later case we can put a lower bound on the kink angle by requiring that the polar angle of the south pole is always greater than that of the north (see Figure 5.15). From our MPE this gives a lower bound of 4.6° for the polar kink angle. This latitudinal kink can be compared with the longitudinal kink of interpulses observed in other pulsars: often these are not found at exactly 180° , but can deviate by 10’s of degrees (see the separation of interpulses for double-pole interpulses in Table 1 of [115]). Allowing for such kinks in both beams is a possible extension to the precessional model.

[†]The use of ‘interpulse’ here strictly refers to seeing the opposite beam of the pulsar, and not cases where the pulsar is almost aligned and interpulses are thought to come from the same beam

[‡]Numbers generated from the maximum posterior estimates of χ and θ using the spin-down data. The estimated error for both values is $\pm 0.08^\circ$

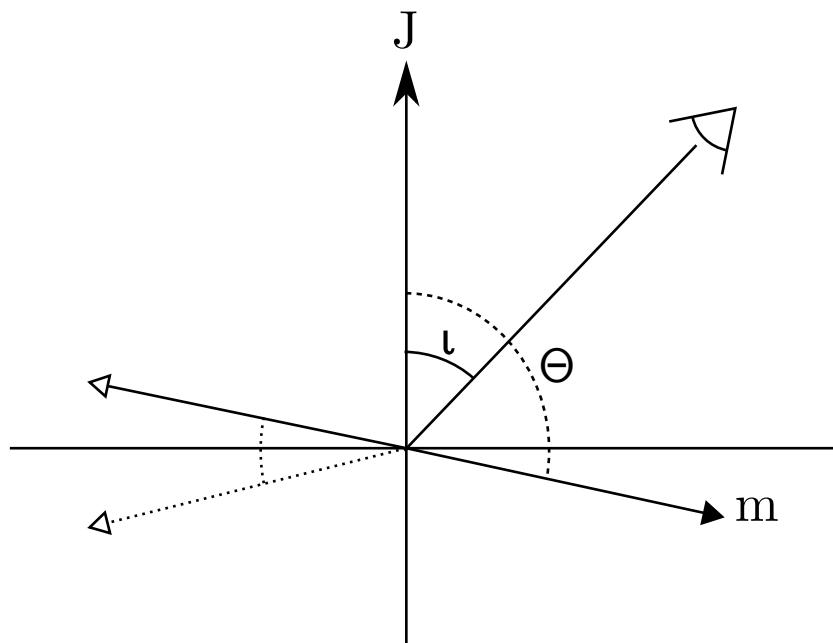


Figure 5.15: Extension to Figure 5.8 adding in the south pole (marked by a white triangle and solid line) for an instance in which the north pole is in the southern hemisphere. The dotted line with a white arrowhead is the ‘kinked’ beam. For clarity, the north and south poles are shown at times that differ by half a rotation period.

Chapter 6

Continuous gravitational waves: calculating the mismatch

Timing variations, in the form of glitches or timing noise, may pose a serious risk to efforts to discover continuous GWs from neutron stars. This is because searches rely on *matched filtering*, which searches the data using a template; if the template and signal don't match, due to timing variations in the signal, the template may be unable to detect the signal. In this chapter, we will give a general introduction to the detection methods used in GW data analysis and then develop our tools which will be used in later chapters to quantify the significance of timing variations for continuous GW searches. The details of estimating the risk posed by glitches to GW searches will be discussed in Chapter 7 and then in Chapter 8 and Chapter 9 we will investigate the role of timing noise for GW searches.

6.1 Introduction

Rotating neutron stars capable of supporting non-axisymmetric mass distributions will emit continuous gravitational waves* (GWs) due to their time-varying quadrupole moments. These may be detectable by ground-based detectors. The emitted signals can persist for much longer than the duration of a typical search, but are weak in amplitude, making them difficult to detect in the noise of the detector.

To find a signal, continuous GW searches use matched filtering techniques such as the \mathcal{F} -statistic [85], which compare the output of the detector with a template. These techniques are powerful, provided that the signal and template remain coherent for the duration of the observation. If the signal can be perfectly matched by a template, then the signal to noise ratio, ρ , used to quantify the detection likelihood, scales as $\rho^2 \propto T_{\text{coh}}$

*Note that in general ‘continuous waves’ can refer to any quasi-monochromatic long-lasting gravitational-wave signals, such as emitted by binaries of white dwarfs, neutron stars or black holes, which would be detectable by LISA or pulsar timing arrays. Here we refer to continuous GWs exclusively in the context of spinning nonaxisymmetric neutron stars as relevant to ground-based detectors.

(e.g. see [140]) where T_{coh} is the coherent observation time. This suggests searching over longer observations increases the chances of making a detection.

The templates must model the monotonic spin-down of the source due to the electromagnetic (EM) and gravitational torque; this is done by Taylor expanding the phase (see Eqn. (1.2)) about some reference time t_{ref} .

Note that all times refer to the solar system barycentre and we assume the timing model has already correctly accounted for the dispersion measure, proper motion and other parameters as discussed in Edwards et al. [57]. Pulsar astronomers fit this model to observed time of arrivals (TOAs). If the best fit model is accurate enough to track the pulsar to within a single rotation, the resulting timing solution is described as *phase-connected*. Often such solutions are capable of tracking the pulsar over durations greater than a year. For gravitational-wave searches, this level of accuracy motivates the use of the same Taylor expansion phase models to account for the spin-down. Pulsar astronomers measure the frequency ν and higher order coefficients describing the rotation of the pulsar itself. In this work we will denote the GW frequency and its derivatives by f, \dot{f} , etc. The relation between the GW frequency and rotation frequency depends on the emission mechanism; in this work we will consider only searches for emission from non-axisymmetric neutron stars such that $f = 2\nu$ (see for example §16.6 of Shapiro and Teukolsky [154]).

The rest of this chapter is organised as follows: in Section 6.2 we will define the mismatch which quantifies the loss of signal to noise ratio for imperfectly matched signals. In Section 6.3 we will familiarise the reader with methods to calculate the mismatch when the phase of the GW is not a smooth Taylor expansion and describe why this method of calculating the mismatch becomes intractable for complicated signals. Then, in Section 6.4 we will define a new approach to calculating the mismatch by splitting the signal into separate subdomains.

6.2 Introduction to the mismatch

While Taylor expansion models are on average reliable enough to track the spin-down, pulsars do show timing variations either in the form of glitches, occasional sudden increases in the rotation frequency, or continuous low-frequency variations known as *timing noise* (see Chapter 2 for an overview). To quantify the effect these variations will cause, we will use the *mismatch*, which is the loss of signal to noise ratio due to an imperfect match between the template and signal. The mismatch is non-zero whenever the signal and template are not perfectly matched. Even if the signal does not contain any timing variations, a mismatch can occur when the template, chosen from a finite grid of points in parameter space, does not share exactly the same parameters as the signal. This problem has been studied in detail since finite computing resources limit the number of templates that can be used in a search, ensuring that any prospective search must have some level of mismatch. In this section, we introduce the tools that have been developed to understand this problem. We will return to the problem of timing variations later on in this chapter.

6.2.1 Defining the mismatch

We begin by assuming that the gravitational-wave detector strain data contains a periodic GW signal with phase $\Phi_s(t; \lambda_s^\alpha)$, where λ_s^α is a vector of the signal parameters. A fully-coherent search consists of applying a matched filtering algorithm over a coherence time T_{coh} to search the data using a *template*; let us then denote the phase evolution of the template as $\Phi_t(t; \lambda_t^\alpha)$, such that λ_t^α is a vector of the template parameters.

Defining the phase-difference between the signal and the template as

$$\Delta\Phi(t; \lambda_t^\alpha, \lambda_s^\alpha) = \Phi_s(t; \lambda_s^\alpha) - \Phi_t(t; \lambda_t^\alpha), \quad (6.1)$$

then, following the work of Prix and Itoh [141] and neglecting amplitude modulations in the signal, we can define the matched filtering amplitude

$$X = \frac{1}{T_{\text{coh}}} \int_{T_{\text{coh}}} e^{i\Delta\Phi} dt, \quad (6.2)$$

This amplitude will have a global maximum of $|X| = 1$ for a perfectly matched signal and will decrease for increasing phase mismatches. In this way, if we define $\tilde{\rho}$ as the signal to noise ratio (SNR) measured in a fully-coherent search, then X defines the loss of SNR incurred due to the non-zero phase difference $\Delta\Phi$ as compared to a perfectly matched signal with SNR $\tilde{\rho}_{\text{pm}}$ such that

$$\tilde{\rho} = |X| \tilde{\rho}_{\text{pm}} \quad (6.3)$$

A simple, dimensionless measure of the loss of SNR is found by rearranging and defining the fully-coherent *mismatch* as

$$\tilde{\mu}(\lambda_s^\alpha, \lambda_t^\alpha) = 1 - |X(\lambda_s^\alpha, \lambda_t^\alpha)|^2, \quad (6.4)$$

such that

$$\tilde{\mu} = \frac{\tilde{\rho}_{\text{pm}}^2 - \tilde{\rho}^2}{\tilde{\rho}_{\text{pm}}^2}. \quad (6.5)$$

6.2.2 Interpreting the mismatch

In this work we will quantify the effect of glitches by the mismatch. However, it may be useful to interpret a mismatch in the following way. For a perfectly matched continuous GW signal it can be shown [85] that the SNR scales as

$$\rho_{\text{pm}} \propto \frac{h_0}{\sqrt{S_n}} \sqrt{T_{\text{coh}} \mathcal{N}}, \quad (6.6)$$

where h_0 is the strain-amplitude for an optimally oriented source with respect to the detector, S_n , the noise measured in the detector and \mathcal{N} is the number of detectors.

Combining this equation with Eqn. (6.5), the SNR of a signal with fully-coherent mismatch $\tilde{\mu}$ scales as

$$\rho \propto \sqrt{1 - \tilde{\mu}} \frac{h_0}{\sqrt{S_n}} \sqrt{T_{\text{coh}} \mathcal{N}}, \quad (6.7)$$

which is to say it reduces the sensitivity by a factor $\sqrt{1 - \tilde{\mu}}$.

6.2.3 Taylor expansion signals and templates

So far we have not yet defined a phase evolution model for either the signal or the template. The signal phase $\Phi_s(t; \lambda_s^\alpha)$ will depend on the GW production mechanism, but assuming the signal is generated by the canonical non-axisymmetric distortion, the GW signal will be produced with a frequency twice that of the rotational frequency and this frequency will have a spin-down rate twice the rotational spin-down rate. From the success of pulsar astronomy, we know that the phase evolution of the rotation can be modelled by a Taylor expansion. For now, let us truncate at the first order in frequency derivative, but we could use higher order terms if required. Therefore, we can begin by modelling the GW phase evolution using a Taylor expansion

$$\Phi_s(t, \lambda_s^\alpha) = \phi_s + 2\pi \left[f_s(t - t_{\text{ref}}) + \frac{\dot{f}_s}{2}(t - t_{\text{ref}})^2 \right], \quad (6.8)$$

such that $\lambda_s^\alpha = [\phi_s, f_s, \dot{f}_s]$. In principle, we can include higher order terms, but for this work we truncate at the first derivative of the frequency.

For blind searches, where we have no hints about the phase evolution, typical continuous GW searches assume that the GW signal is a Taylor expansion and search for it using a Taylor expansion template. Defining $\lambda_t^\alpha = [\phi_t, f_t, \dot{f}_t]$ as the Taylor expansion template parameters, Eqn. (6.1), the phase difference, is

$$\Delta\Phi(t; \lambda_s^\alpha, \lambda_t^\alpha) = \Delta\phi + 2\pi \left[\Delta f(t - t_{\text{ref}}) + \frac{\Delta \dot{f}}{2}(t - t_{\text{ref}})^2 \right], \quad (6.9)$$

where $\Delta\lambda^\alpha = \lambda_s^\alpha - \lambda_t^\alpha$ is the difference between the signal and the template which we refer to as the *parameter space offset*. Note that by choosing any two of $\lambda_s^\alpha, \lambda_t^\alpha$, and $\Delta\lambda^\alpha$ we have a choice of three equivalent parameterisations. When the parameter offset $\Delta\lambda^\alpha$ vanishes, the matched filtering amplitude tends to unity and hence the mismatch tends to zero. In the other extreme, for a suitably large parameter space offsets, the mismatch approaches zero and so the signal is completely lost.

Substituting Eqn. (6.9) into Eqn. (6.4) gives the mismatch between a smooth Taylor expansion signal and template. In this work we will calculate the mismatch for a signal that is not a smooth Taylor expansion. This can be done in this way for some special cases. For example, taking $\Delta\lambda^\alpha = [\Delta\phi, 0, 0]$, the mismatch is found to be zero: this reflects the fact that fully-coherent matched filtering is insensitive to any arbitrary phase offset between the signal and template. In more complicated cases, the integrals become intractable. To aid in calculation, we will use the so-called metric approximation. In

the next section we will introduce it in the context of smooth Taylor expansion signals. Later on in Section 6.4 we will describe how we extend it to arbitrary signals.

6.2.4 The metric-mismatch approximation for fully-coherent searches

The fully-coherent mismatch Eqn. (6.4) has a local minimum of zero at $\Delta\lambda^\alpha = 0$. Expanding about this minimum up to the leading order term, Brady et al. [35] approximated the mismatch by

$$\tilde{\mu}(\lambda_t^\alpha, \Delta\lambda^\alpha) \approx g_{\alpha\beta} \Delta\lambda^\alpha \Delta\lambda^\beta, \quad (6.10)$$

where $g_{\alpha\beta}$ is the parameter space metric given by

$$g_{\alpha\beta} = \frac{1}{2} \partial_\alpha \partial_\beta \tilde{\mu}(\lambda_t^\alpha, \Delta\lambda^\alpha) \Big|_{\Delta\lambda^\alpha=0}. \quad (6.11)$$

Note that we define $\partial_\alpha \equiv \partial_{\Delta\lambda^\alpha}$. The metric is a function of both the total coherence time T_{coh} over which the matched filter is performed and the reference time at which the Taylor expansions are defined, but not the signal itself. Partially evaluating the metric, we find that

$$\begin{aligned} g_{\alpha\beta} = & \frac{1}{T_{\text{coh}}} \int_0^{T_{\text{coh}}} \partial_\alpha \Delta\Phi \partial_\beta \Delta\Phi dt \\ & - \frac{1}{T_{\text{coh}}^2} \int_0^{T_{\text{coh}}} \partial_\alpha \Delta\Phi dt \int_0^{T_{\text{coh}}} \partial_\beta \Delta\Phi dt. \end{aligned} \quad (6.12)$$

This metric formulation provides a method to measure the *metric-mismatch* between a signal and template that are both Taylor expansions. Brady et al. [35] proposed this method with the aim of picking the spacing of templates in parameter space such that the maximum allowable mismatch would not rise above a pre-defined threshold. In this work, we will instead use this metric-mismatch approximation to calculate mismatches between signals and Taylor expansion templates.

It is worth commenting that the full mismatch, as calculated from Eqn. (6.4), is bounded by $[0, 1]$. In contrast, the approximate metric-mismatch is bounded by $[0, \infty)$. This is because the expansion of the mismatch in Eqn. (6.10) was taken about $\Delta\lambda^\alpha = 0$ and so for sufficiently large parameter space offsets the expansion breaks down. In this case, the metric-mismatch approximation non-linearly overestimates the true mismatch. A metric-mismatch above one, while losing the interpretation as a direct loss of SNR, still corresponds to a large true mismatch, and hence a significant loss of SNR.

6.2.5 The mismatch for semi-coherent searches

Fully-coherent searches are computationally demanding, so much so that it is infeasible to perform fully-coherent searches for wide-parameter searches such as the all-sky search. Instead, *semi-coherent* searches are used which require far fewer templates and result in

more sensitive searches at a fixed computing cost [140]. There are numerous implementations of semi-coherent searches, for example the E@H search uses the Hough-transform method [99]; in this section we will discuss the generic case.

A semi-coherent search divides an observation time T_{obs} into N_{seg} segments of duration T_{coh} . Each segment is searched fully coherently with a resulting SNR $\tilde{\rho}_j^2$ where j labels the segment. The semi-coherent search then recombines these segments by summing all segments at the same point in parameter space λ_t^α to give a new detection statistic

$$\hat{\rho}^2(\lambda_t^\alpha, \Delta\lambda^\alpha) = \sum_j^{N_{\text{seg}}} \tilde{\rho}_j^2(\lambda_t^\alpha, \Delta\lambda^\alpha), \quad (6.13)$$

where the ‘hat’ denotes that it is a semi-coherent quantity.

For each fully-coherent segment, we can rearrange Eqn. (6.5) to give

$$\tilde{\rho}_j^2(\lambda_t^\alpha, \Delta\lambda^\alpha) = \tilde{\rho}_{\text{pm}}^2(1 - \tilde{\mu}_j(\lambda_t^\alpha, \Delta\lambda^\alpha)). \quad (6.14)$$

Here we take $\tilde{\rho}_{\text{pm}}$ to be the same for all segments such that it does not carry a j -index; in this way we neglect variations in the signal such as the signal amplitude or the motion of the Earth and also require all segments to be of equal duration. For the semi-coherent detection statistic we may define $\hat{\rho}_{\text{pm}}^2$ as the squared SNR in the absence of any mismatch. Then the semi-coherent mismatch $\hat{\mu}$ can be implicitly defined as

$$\hat{\rho}^2(\lambda_t^\alpha, \Delta\lambda^\alpha) = \hat{\rho}_{\text{pm}}^2(1 - \hat{\mu}(\lambda_t^\alpha, \Delta\lambda^\alpha)), \quad (6.15)$$

where $\hat{\mu}$ is the semi-coherent mismatch. When there is no mismatch, the sum of the $\tilde{\rho}_{\text{pm}}^2$ is equal the semi-coherent $\hat{\rho}_{\text{pm}}^2$ such that

$$\hat{\rho}_{\text{pm}}^2 = N_{\text{seg}} \tilde{\rho}_{\text{pm}}^2. \quad (6.16)$$

Inserting all these expressions into Eqn. (6.13) we see that the semi-coherent mismatch is an average over the individual full-coherent segment mismatches:

$$\hat{\mu}(\lambda_t^\alpha, \Delta\lambda^\alpha) = \frac{1}{N_{\text{seg}}} \sum_{j=1}^{N_{\text{seg}}} \tilde{\mu}_j(\lambda_t^\alpha, \Delta\lambda^\alpha). \quad (6.17)$$

With this expression, we can first calculate the mismatch for each fully-coherent segment, and then calculate the mismatch for a semi-coherent search. However, we must be careful to sum all the mismatches at the same point in parameter space.

In this section we have introduced: the fully-coherent mismatch, used to quantify the loss of SNR in fully-coherent GW searches; the metric-mismatch approximation, a useful way to calculate mismatches that avoids the complicated integration; and the semi-coherent mismatch, quantifying the loss of SNR in semi-coherent GW searches. The mismatch is a useful way to quantify how well a search performs given a particular GW signal. However, the tools discussed in this section to calculate this mismatch apply only to smooth Taylor expansion signals and templates. Signals which contain timing

variations are by definition not smooth Taylor expansions. Therefore, we need a method to calculate the mismatch for arbitrary signals. In the next section, we will discuss some special cases of timing variations where the mismatch can be calculated exactly. Then in the final section, we will introduce a new tool which can be used to calculate the mismatch for arbitrary signals.

6.3 Exact mismatch from irregularities in the phase

The \mathcal{F} -statistic considered in Brady et al. [35] analytically minimises over the phase. Therefore, if the signal and template can both be modelled by a single smooth Taylor expansion, but with a finite phase-offset, the mismatch is zero: the mismatch is insensitive to any overall phase difference between the signal and template. However, a mismatch will occur if the overall phase between the signal and template changes during an observation. In this section, we will investigate three scenarios where this occurs. While not all these scenarios are physically relevant to real astrophysical systems, this introduces some concepts in a simple setting before we tackle the more difficult real astrophysical systems.

6.3.1 Two subdomains with a phase discontinuity

We begin with a simple system in which the signal undergoes an instantaneous ‘jump’ in its GW phase halfway through the observation. We can model this signal by a *piecewise* Taylor expansion with two subdomains. Both subdomains are of equal duration and follow a smooth spin-down, except that there is a phase discontinuity at their interface; we illustrate this setup in Figure 6.1. Parameterising by the offset with respect to the

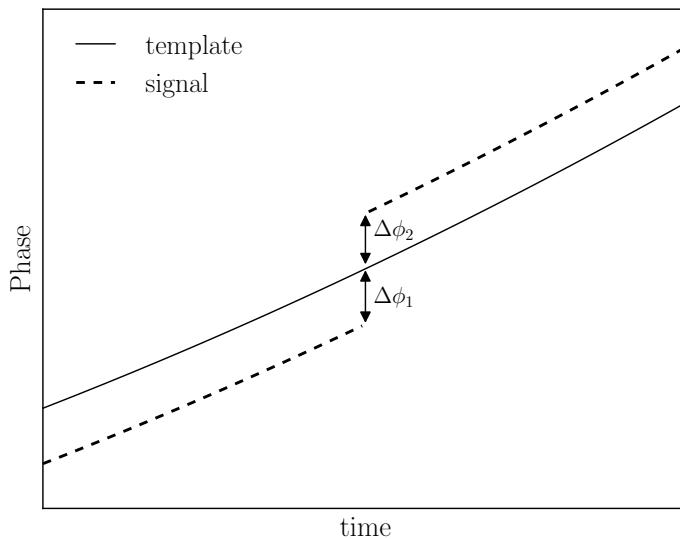


Figure 6.1: Illustration of the signal and template defined in equation (6.18)

template parameters λ_0 for some arbitrary template and phase jump, we write the phase deviations in the two subdomains as

$$\Delta\Phi(t) = \begin{cases} \Delta\phi_1 & 0 < t < T/2 \\ \Delta\phi_2 & T/2 < t < T \end{cases}. \quad (6.18)$$

To compute the matched filtering amplitude given in Eqn. (6.2), we can factorise the integral using the additivity of integration on intervals into two integrations

$$X = \frac{1}{T} \int_0^T e^{i\Delta\Phi(t)} dt \quad (6.19)$$

$$= \frac{1}{T} \left(\int_0^{T/2} e^{i\Delta\phi_1} dt + \int_{T/2}^T e^{i\Delta\phi_2} dt \right) \quad (6.20)$$

$$= \frac{1}{2} (e^{i\Delta\phi_1} + e^{i\Delta\phi_2}). \quad (6.21)$$

Because our choice in splitting up the integral exactly matches the subdomains defined in Eqn. (6.18) and in each subdomain $\Delta\Phi(t)$ is constant, we can compute the integrals. The absolute square value of the matched filtering amplitude is then

$$|X|^2 = \frac{1}{4} (e^{i\Delta\phi_1} + e^{i\Delta\phi_2}) (e^{-i\Delta\phi_1} + e^{-i\Delta\phi_2}) \quad (6.22)$$

$$= \frac{1}{4} (2 + e^{i(\Delta\phi_1 - \Delta\phi_2)} + e^{-i(\Delta\phi_1 - \Delta\phi_2)}) \quad (6.23)$$

$$= \frac{1}{2} (1 + \cos(\Delta\phi_1 - \Delta\phi_2)). \quad (6.24)$$

Finally, the fully-coherent mismatch can be calculated from Eqn. (6.4) to be

$$\tilde{\mu} = \frac{1}{2} (1 - \cos(\Delta\phi_1 - \Delta\phi_2)). \quad (6.25)$$

From this result, we learn that it is not the phase difference with respect to each of the Taylor expansions ($\Delta\phi_1$ or $\Delta\phi_2$) that is important, but the total phase jump at their interface. For $\Delta\phi_1 = \Delta\phi_2$ the mismatch vanishes, this recovering the case of a single Taylor expansion signal with an arbitrary overall phase offset, which always has zero mismatch. This result can be used to calculate the mismatch due to a glitch, if the glitch is purely in the phase. In reality, we know that glitches are discontinuities in the frequency and spin-down rate, this will be discussed later in Chapter 7, but this simple calculation provides a stepping stone to understanding the more complicated complete result.

6.3.2 N subdomains with phase discontinuities

We can further generalise Eqn. (6.25) by letting the signal be comprised of N equal duration subdomains with a phase discontinuity $\Delta\phi_i$ for the i^{th} subdomain. Then, the

matched filtering amplitude can be written

$$X = \frac{1}{T} \left(\int_0^{t_1} e^{i\Delta\phi_1} dt + \int_{t_1}^{t_2} e^{i\Delta\phi_2} dt + \cdots + \int_{t_{N-1}}^{t_N} e^{i\Delta\phi_N} dt \right) \quad (6.26)$$

$$= \frac{1}{T} \left(\frac{T}{N} e^{i\Delta\phi_1} + \frac{T}{N} e^{i\Delta\phi_2} + \cdots + \frac{T}{N} e^{i\Delta\phi_N} \right) \quad (6.27)$$

$$= \frac{1}{N} \sum_{i=1}^N e^{i\Delta\phi_i}. \quad (6.28)$$

Squaring the matched filtering amplitude and simplifying

$$|X|^2 = \frac{1}{N^2} \left(\sum_{i=1}^N e^{i\Delta\phi_i} \right) \left(\sum_{j=1}^N e^{-i\Delta\phi_j} \right) \quad (6.29)$$

$$= \frac{1}{N^2} \sum_{i=1}^N e^{i\Delta\phi_i} \left(\sum_{j=1}^N e^{-i\Delta\phi_j} \right) \quad (6.30)$$

$$= \frac{1}{N^2} \sum_{i=1}^N e^{i\Delta\phi_i} \left(e^{-i\Delta\phi_i} + \sum_{\substack{j=1 \\ j \neq i}}^N e^{-i\Delta\phi_j} \right) \quad (6.31)$$

$$= \frac{1}{N^2} \left(\sum_{i=1}^N 1 + \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N e^{i(\Delta\phi_i - \Delta\phi_j)} \right) \quad (6.32)$$

$$= \frac{1}{N^2} \left(N + \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \cos(\Delta\phi_i - \Delta\phi_j) + i \sin(\Delta\phi_i - \Delta\phi_j) \right). \quad (6.33)$$

In this final summation, for each pair (i, j) the corresponding pair (j, i) will exist in the sum. This leads to a cancellation of the imaginary part and a doubling of the real part

$$|X|^2 = \left(\frac{1}{N} + \frac{1}{N^2} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \cos(\Delta\phi_i - \Delta\phi_j) \right). \quad (6.34)$$

Then the fully-coherent mismatch is given by:

$$\tilde{\mu} = 1 - \frac{1}{N} - \frac{1}{N^2} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \cos(\Delta\phi_i - \Delta\phi_j). \quad (6.35)$$

This result could be used to estimate the mismatch due to a random walk model of timing noise where the random walk was purely in the phase (see Section 2.3.1 for example). In Chapter 9 we will discuss random walks in GW signals in more detail and also calculate the mismatch for random walks in the frequency and spin-down rate.

6.3.3 Oscillating phase deviations

Timing residuals, or equivalently the phase offsets between Taylor expansions and the real signals, are often found to have a quasi-periodic structure [81]. It is reasonable to ask if, because the residuals oscillate about the origin, the loss of detection will be cancelled out. In this section we consider such an oscillating phase deviation and show that, over many oscillations, the mismatch is non-zero.

We model a small sinusoidal phase offset by a trigonometric function

$$\Delta\Phi(t) = \varepsilon \cos(\omega t), \quad (6.36)$$

where ε is the amplitude of variations and ω is the angular frequency of the oscillations. If we assume that the variations are small $\varepsilon \ll 1$, we can Taylor expand the exponential phase offset in the matched filtering amplitude

$$X = \frac{1}{T} \int_T 1 + i\varepsilon \cos(\omega t) - \frac{\varepsilon^2}{2} \cos^2(\omega t) - \frac{i\varepsilon^3}{6} \cos^3(\omega t) dt + \mathcal{O}(\varepsilon^4). \quad (6.37)$$

Inserting this into Eqn. (6.4) and computing the fully-coherent mismatch we find that

$$\tilde{\mu} = \varepsilon^2 \left(\frac{1}{2} + \frac{\sin(2T\omega)}{4T\omega} + \frac{\sin^2(T\omega)}{T^2\omega^2} \right) + \mathcal{O}(\varepsilon^4). \quad (6.38)$$

In the limit $T \rightarrow \infty$, this tends to $\varepsilon^2/2$. The mismatch oscillates around this value over observation times similar to the oscillation period; this is shown in Figure 6.2.

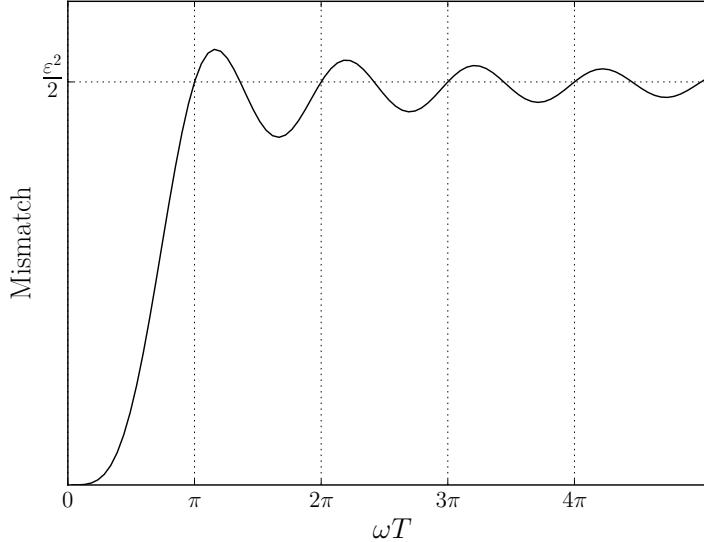


Figure 6.2: Plot demonstrating the behaviour of Eqn. (6.38), the mismatch for an oscillating phase offset.

This simple model demonstrates that if the signal and template are on average coherent, but the phase residual between them oscillates about the mean, then over several cycles

the mismatch will approach a constant value, which depends on the maximum amplitude of the residuals. For a quasi-periodic signal, we expect the same overall picture will emerge with the size of the mismatch dependent on the mean average size of oscillations in the residual.

6.4 Generalising the metric-mismatch approximation to arbitrary signals

In the previous section, we discussed some special cases where the mismatch can be calculated exactly for a signal which has a phase which can't be described by a single Taylor expansion. For more general signals, where the frequency and spin-down rate may vary, the equations soon become unwieldy. In Section 6.2.4 we introduced the metric-mismatch, a tool to estimate the loss of SNR for Taylor expansion signals and templates in fully-coherent searches. In this section, we will extend the fully-coherent metric-mismatch to handle cases where the signal is not a Taylor expansion, but can be approximated by a *piecewise* Taylor expansion; this will allow us to handle any arbitrary phase evolution of the signal. By piecewise Taylor expansion we mean breaking any arbitrary signal into N_{sd} *subdomains* each of which we describe the signal by a single Taylor expansion; the resulting collection of all individual Taylor expansions is a piecewise Taylor expansion. We will distinguish between a *subdomain*, which describes a part of the piecewise function and a *segment*, which is the division of an observation period for a semi-coherent search as discussed in Section 6.2.5.

This *generalised metric-mismatch* will be used in Chapter 7 to calculate the mismatch due to glitches and in Chapter 9 to calculate the mismatch due to a random walk model of timing noise.

6.4.1 Piecewise Taylor expansion

Before we get into the details of formulating a general metric-mismatch, it is worth stating some details about piecewise Taylor expansions. Allowing the signal to be piecewise amounts to placing a second index on the signal parameters, $\lambda_s^{\alpha a}$, which labels the subdomain of the piecewise function. Note that we discriminate between Greek indices, labelling the parameter components, and Roman indices which label the subdomain. For the a^{th} subdomain of the piecewise function, we define t_{ref}^a as the reference time of the Taylor expansion in that subdomain. These can be defined arbitrarily, although it is usual to either define them relative to each subdomain, or at a fixed value for all the subdomains. The piecewise signal function in the a^{th} subdomain is then

$$\Phi_s(t; \{\lambda_s^{\alpha a}\}) = \phi_s^a + 2\pi \left(f_s^a(t - t_{\text{ref}}^a) + \frac{\dot{f}_s^a}{2}(t - t_{\text{ref}}^a)^2 \right), \quad (6.39)$$

when $t_a < t \leq t_{a+1}$ and otherwise undefined: in future this condition will be assumed. By $\{\lambda_s^{\alpha a}\}$ we indicate the set of all signal parameters and by implication their reference times.

Since we will be calculating the phase difference between the signal and template it is convenient to similarly make the template a piecewise function with parameters $\lambda_t^{\alpha a}$ such that

$$\Phi_t(t; \{\lambda_t^{\alpha a}\}) = \phi^a + 2\pi \left(f_t^a(t - t_{\text{ref}}^a) + \frac{\dot{f}_t^a}{2}(t - t_{\text{ref}}^a)^2 \right). \quad (6.40)$$

Then we may write the parameter space difference as $\Delta\lambda^{\alpha a} = \lambda_s^{\alpha a} - \lambda_t^{\alpha a}$, where both the signal and template must refer to the same reference time.

Since we have set up the template phase in Eqn. (6.40) as a piecewise Taylor expansion, the template can be any arbitrary function described by a piecewise Taylor expansion. However, all current and planned searches use a single global Taylor expansion template. To model this, we must introduce *consistency relations* which ensure that the local subdomains all lie along a single global template with parameters $[\phi_t, f_t, \dot{f}_t]$ defined at t_{ref} . These consistency relations may be written as

$$\begin{aligned} \phi_t^a &= \phi_t + f_t(t_{\text{ref}}^a - t_{\text{ref}}) + \frac{\dot{f}_t^2}{2}(t_{\text{ref}}^a - t_{\text{ref}})^2, \\ f_t^a &= f_t + \dot{f}_t(t_{\text{ref}}^a - t_{\text{ref}}), \\ \dot{f}_t^a &= \dot{f}_t. \end{aligned} \quad (6.41)$$

Note the subtle distinction between t_{ref} , the global reference time and t_{ref}^a the reference time for the a^{th} subdomain.

Finally, we can generalise Eqn. (6.9) for the phase difference to the piecewise Taylor expansion description of the signal and template:

$$\Delta\Phi^a(t) = \Phi_s(t; \lambda_s^{\alpha a}) - \Phi_t(t; \lambda_t^{\alpha a}) \quad (6.42)$$

$$= \Delta\phi^a + 2\pi \left(\Delta f^a(t - t_{\text{ref}}^a) + \frac{\Delta \dot{f}^a}{2}(t - t_{\text{ref}}^a)^2 \right). \quad (6.43)$$

6.4.2 The generalised metric-mismatch approximation for fully-coherent searches

We will now generalise the metric-mismatch approximation first introduced by Brady et al. [35] to the case where the signal and template are piecewise Taylor expansions. This calculation follows that given in Section 6.2.4 with the addition of an index labelling the subdomains.

Using a piecewise Taylor expansion to describe the signal and template, the discrete nature of the phase offset $\Delta\Phi^a$ allows us to partition the matched filtering amplitude due to the additivity of integration on intervals. The matched filtering amplitude, given

the set of template parameters and parameter offsets $\{\lambda_t^{\alpha a}, \Delta\lambda^{\alpha a}\}$, is then

$$X(\{\lambda_t^{\alpha a}, \Delta\lambda^{\alpha a}\}) = \frac{1}{T_{\text{coh}}} \int_0^{T_{\text{coh}}} e^{i\Delta\Phi(t)} dt \quad (6.44)$$

$$= \frac{1}{T_{\text{coh}}} \sum_c^{N_{\text{sd}}} \int_{t^c} e^{i\Delta\Phi^c} dt, \quad (6.45)$$

where the integration is taken over the bounds of the c^{th} subdomain. The mismatch, in this general formulation, is then defined by

$$\tilde{\mu}(\{\lambda_t^{\alpha a}, \Delta\lambda^{\alpha a}\}) = 1 - |X(\{\lambda_t^{\alpha a}, \Delta\lambda^{\alpha a}\})|^2 \quad (6.46)$$

The mismatch has a local minimum of zero when $\{\Delta\lambda^{\alpha a}\} = 0$. Expanding in powers of $\Delta\lambda^{\alpha a}$, the leading order term is

$$\tilde{\mu}(\{\lambda_t^{\alpha a}, \Delta\lambda^{\alpha a}\}) = \frac{1}{2} \partial_{\alpha a} \partial_{\beta b} \tilde{\mu}(\{\lambda_t^{\alpha a}, \Delta\lambda^{\alpha a}\}) \Delta\lambda^{\alpha a} \Delta\lambda^{\beta b}, \quad (6.47)$$

where $\partial_{\alpha a} \equiv \partial_{\Delta\lambda^{\alpha a}}$. The metric in this formalism is identified as

$$g_{\alpha\beta ab} = \frac{1}{2} \partial_{\alpha a} \partial_{\beta b} \tilde{\mu}(\{\lambda_t^{\alpha a}, \Delta\lambda^{\alpha a}\}) \Big|_{\{\Delta\lambda^{\alpha a}\}=0}. \quad (6.48)$$

We can then define the fully-coherent *generalised metric-mismatch* as

$$\tilde{\mu}(\{\lambda_t^{\alpha a}, \Delta\lambda^{\alpha a}\}) = g_{\alpha\beta ab} \Delta\lambda^{\alpha a} \Delta\lambda^{\beta b}. \quad (6.49)$$

On the right hand side we sum over the repeated indices. The mismatch is a scalar value quantifying the loss of signal to noise due to the set of parameter offsets $\{\Delta\lambda^{\alpha a}\}$.

Partially evaluating the metric we have

$$\begin{aligned} g_{\alpha\beta ab} &= \frac{1}{T_{\text{coh}}} \sum_c \int_{t_c} \partial_{\alpha a} \Delta\Phi^c \partial_{\beta b} \Delta\Phi^c dt \\ &\quad - \frac{1}{T_{\text{coh}}^2} \sum_b \int_{t_c} \partial_{\alpha a} \Delta\Phi^c dt \sum_{b'} \int_{t_{c'}} \partial_{\beta c'} \Delta\Phi^{c'} dt. \end{aligned} \quad (6.50)$$

Expanding the summation over the a index, we note that,

$$\partial_{\alpha a} \Delta\Phi^c = [1, 2\pi(t - t_r^a), \pi(t - t_r^a)^2]^\alpha \delta_{ac}, \quad (6.51)$$

and so we can simplify Eqn. (6.50) by dropping the terms which vanish from the summation

$$\begin{aligned} g_{\alpha\beta ab} &= \frac{1}{T_{\text{coh}}} \delta_{ab} \int_{t_a} \partial_{\alpha a} \Delta\Phi^a \partial_{\beta a} \Delta\Phi^a dt \\ &\quad - \frac{1}{T_{\text{coh}}^2} \int_{t_a} \partial_{\alpha a} \Delta\Phi^a dt \int_{t_b} \partial_{\beta b} \Delta\Phi^b dt. \end{aligned} \quad (6.52)$$

This expression, given a choice of decomposition of the observation time T into N_{sd} (not

necessarily equal) time subdomains labelled t_a , will result in a rank 4 tensor. Then, given a set of $\{\Delta\lambda^{\alpha a}\}$ which together with the decomposition of the time into subdomains defines the signal and template evolution, the fully-coherent metric-mismatch can be calculated from Eqn. (6.49).

6.4.3 Explicit calculation of the metric

In this final subsection, we will derive two explicit calculation of the metric given choices for the reference times. These will be used later on in this thesis.

6.4.3.1 Reference times in the middle of the subdomains

If there are N_{sd} subdomains with the a^{th} subdomain of duration ΔT_a then we can bound the integration over each subdomain by $t_a, t_a + \Delta T_a$. We are free to choose the reference time in any way we like; for this example, we define the reference time in each subdomain to be halfway through, that is $t_{\text{ref}}^a = t_a + \Delta T_a/2$. Using the following identity

$$\int_{t_a}^{t_a + \Delta T_a} \left(t - \left(t_a + \frac{\Delta T_a}{2} \right) \right)^n dt \equiv \frac{(1 + (-1)^n)}{n+1} \left(\frac{\Delta T_a}{2} \right)^{n+1}, \quad (6.53)$$

we have that

$$\int_{t_a}^{t_a + \Delta T_a} \partial_\alpha \Delta \Phi^a \partial_\beta \Delta \Phi^a dt = \begin{bmatrix} \Delta T_a & 0 & \frac{\pi \Delta T_a^3}{12} \\ 0 & \frac{\pi^2 \Delta T_a^3}{3} & 0 \\ \frac{\pi \Delta T_a^3}{12} & 0 & \frac{\pi^2 \Delta T_a^5}{80} \end{bmatrix}^{\alpha\beta}, \quad (6.54)$$

and

$$\int_{t_a}^{t_a + \Delta T_a} \partial_\alpha \Delta \Phi^a dt = \left[\Delta T, 0, \frac{\pi \Delta T_a^3}{12} \right]^\alpha. \quad (6.55)$$

which can be inserted into Eqn. (6.52) to calculate the metric.

The rank 4 metric $g_{\alpha\beta ab}$ is degenerate in the subdomain indices a and b having only two distinct terms. We can distinguish these by either $a = b$ or $a \neq b$. This allows to write the metric $g_{\alpha\beta ab}$ compactly as

$$g_{\alpha\beta ij} = \begin{bmatrix} \delta_{ab} \frac{\Delta T_a}{T_{\text{coh}}} - \left(\frac{\Delta T_a}{T_{\text{coh}}} \right)^2 & 0 & \frac{\delta_{ab}}{T_{\text{coh}}} \frac{\pi \Delta T_a^3}{12} - \frac{\pi \Delta T_a^4}{12 T_{\text{coh}}^2} \\ 0 & \delta_{ab} \frac{\pi^2 \Delta T_a^3}{3 T_{\text{coh}}} & 0 \\ \frac{\delta_{ab}}{T_{\text{coh}}} \frac{\pi \Delta T_a^3}{12} - \frac{\pi \Delta T_a^4}{12 T_{\text{coh}}^2} & 0 & \frac{\delta_{ab}}{T_{\text{coh}}} \frac{\pi^2 \Delta T_a^5}{80} - \left(\frac{\pi \Delta T_a^3}{12 T_{\text{coh}}} \right)^2 \end{bmatrix}^{\alpha\beta}. \quad (6.56)$$

6.4.3.2 Reference times at the start of each subdomain

If instead we set the reference time to be at the start of each subdomain then, following the method used to derive Eqn. (6.56), the metric can be written compactly as

$$g_{\alpha\beta ij} = \begin{bmatrix} \delta_{ij}N^{-1}-N^{-2} & \pi\Delta T(\delta_{ij}N^{-1}-N^{-2}) & \frac{\Delta T^2\pi}{3}(\delta_{ij}N^{-1}-N^{-2}) \\ \pi\Delta T(\delta_{ij}N^{-1}-N^{-2}) & \pi^2\Delta T^2\left(\delta_{ij}\frac{4N^{-1}}{3}-N^{-2}\right) & \pi^2\Delta T^3\left(\delta_{ij}\frac{N^{-1}}{2}-\frac{N^{-2}}{3}\right) \\ \frac{\Delta T^2\pi}{3}(\delta_{ij}N^{-1}-N^{-2}) & \pi^2\Delta T^3\left(\delta_{ij}\frac{N^{-1}}{2}-\frac{N^{-2}}{3}\right) & \pi^2\Delta T^4\left(\frac{N^{-1}}{5}-\frac{N^{-2}}{9}\right) \end{bmatrix}^{\alpha,\beta}. \quad (6.57)$$

Chapter 7

Glitches in continuous gravitational waves

Continuous gravitational waves from neutron stars could provide an invaluable resource to learn about their interior physics. One common detection method involves matched filtering a modelled template against the noisy gravitational wave data to find unknown signals. As introduced in Chapter 6, this method suffers a mismatch (a loss of signal to noise ratio) if the unknown signal deviates from the template. One significant way this may happen is if the neutron star undergoes a glitch, a sudden rapid increase in the rotation frequency as introduced in Section 2.1, a phenomenon seen in the timing of many radio pulsars. While the mechanism which causes pulsars to glitch is not fully understood, it is likely that any continuous gravitational waves emitted by the star will also be affected by the glitch since both are intimately related to the neutron star crust.

In this chapter, we use information on the rate and size of pulsar glitches, as deduced from the observed population of glitching radio pulsars, to estimate the potential mismatch introduced when searching for gravitational waves from neutron stars whose rotational timing is unknown. We intend to publish this chapter, along with the generalised mismatch defined in Sec. 6.4, in due course.

7.1 Introduction

Electromagnetic (EM) observations of pulsar glitches have long been one of the most fruitful sources of insight into neutron star physics. Glitches are characterised by a sudden increase in the rotation frequency, often accompanied by a jump in the frequency derivative and an exponential recovery of some fraction of the initial frequency jump. The events happen rapidly and are sufficiently disruptive that pulsar timing models often loose phase coherence over the event.

Two leading models exist to explain glitches. In the *superfluid pinning* model, some portion of the interior superfluid is pinned, and does not participate in the smooth torque-driven spin-down of the rest of the crust (where ‘crust’ refers to the actual crust,

plus whatever other parts of the star that are strongly coupled to it). After some period, the crust will therefore have developed a frequency lag compared to the pinned superfluid. A glitch occurs when the two components recouple, transferring angular momentum from the pinned superfluid to the crust and producing a spin-up of the crust [16, 15]. Alternatively glitches could be caused by *crust cracking* as the crust readjusts to a minimum energy configuration brought about by the gradual decay of the spin-down rate [28]. It is also possible that glitches result from a combination of these two models; evidence for this was found by Melatos et al. [123]. In either case it seems reasonable to assume that both the crust and the core will be involved.

Rotating isolated neutron stars can produce continuous gravitational-wave (GW) emission from non-axisymmetric distortions, colloquially also known as ‘mountains’. These require the mountain to be supported by either elastic stresses in the crust or magnetic fields. In this model, the star emits a monochromatic GW at a frequency f_s which is twice the rotation frequency f_s , i.e., $f_s = 2\nu$.

It is unclear exactly how a glitch, as observed in the EM channel, will manifest in the GW channel. Observing a glitch in both the EM and GW channels would provide a unique opportunity to investigate the neutron star interior, but to do this we must first detect continuous GW emission.

Estimates for the intrinsic gravitational wave strain amplitude h_0 for canonical models of GW emissions (see for example Abbott et al. [11]) suggest they are extremely weak compared to the noise level of advanced detectors [5]. As a result, in order to detect a signal, significant effort has been put into advanced data analysis methods, which may be capable of identifying the putative signals. Many of these methods rely on *matched filtering* in which a template is correlated with the data with the hope of detecting the presence of the unknown signal similar to the template. The power of these methods is that the signal-to-noise ratio (SNR) scales as the square-root of the observation time [140]. Due to the longevity of GW signals, this allows the weak signal to be discerned from the noise.

These methods are powerful, but harbour a vulnerability in any instance where the template, a Taylor expansion in the phase usually up to second order, does not match the signal. For GW signals from non-axisymmetric distortions of neutron stars, we can expect that discrepancies from such a template may manifest in one of two ways. Firstly, we know from radio pulsar timing that the spin-down of a pulsar differs from a Taylor expansion due to timing noise. We will consider the effect of timing noise in Chapter 8 and Chapter 9; in this chapter we will address the second potential manifestation: glitches.

There are two distinct questions to answer in the case of glitches: ‘how probable is it that a glitch will occur during our GW observation?’, and ‘if a glitch does occur, what effect will it have on our ability to discover the GW signal?’. To answer these questions, we first use known radio pulsar glitch statistics to estimate the size and rate of radio pulsar glitches for the parameter spaces considered in continuous GW searches. Then, we assume that the EM and GW channels are locked in phase, such that a glitch observed in radio pulsations also exist in the GW signal at twice the phase, frequency

and spin-down rate. Having obtained an estimate for the magnitude and rate of glitches for potential continuous GW sources, we quantify the effect such glitches will have on current continuous GW detection methods by calculating the mismatch. We will do this using the generalised metric-mismatch introduced in Section 6.4. Specifically, we model a glitch as a piecewise Taylor expansion with a discontinuity at the glitch; we do not model the exponential recovery observed in some glitches, but we will discuss the significance this may have in Section 7.5.4. Ultimately, the goal of this chapter is to estimate the risk faced by current and ongoing continuous GW searches to glitches in their target population.

The remainder of this chapter is organised in the following way. We begin in Section 7.2, with a description of current GW searches and how glitches may effect them. Then, in Section 7.3 we investigate the statistical properties of the observed radio pulsar glitches. In Section 7.4 we calculate, given the properties of a single glitch, the corresponding mismatch for a semi-coherent and fully-coherent search. To give quantitative results relevant to real GW searches, in Section 7.5 we use these calculations to transform the predicted glitch sizes into predicted mismatches and rates for GW searches.

7.2 Continuous gravitational-wave searches

Searches which target a particular source making use of the EM emission (for example the Abbott et al. [13] targeted search for the Crab pulsar) are able to handle the epoch of a glitch, either by avoiding searches over the glitch, or allowing for a jump in the timing solution at that point [13]. By this merit, such searches have a very low risk of being disrupted by a glitch coupled to the EM channel, provided the GW channel follows closely the phase evolution of the EM channel.

In contrast, blind GW searches which, by definition, search for signals without an EM counterpart do not have any such prior knowledge. This category of searches includes both *directed* searches where a small patch of sky is searched in which a neutron star is believed to exist (e.g. the Aasi et al. [6] search for continuous GWs from supernova remnants), and *all-sky* searches; in both instances a band of frequencies and frequency-derivatives are usually searched since they are inherently unknown. Typically, these searches match filter against smooth templates built from a Taylor expansion in the phase; as such the templates do not include glitches. If a neutron star, emitting detectable levels of GW emission, undergoes a glitch in the GW channel, then the matched filtering method may fail because the template is a poor match to the real, glitching signal.

All-sky and targeted blind searches are most notably at risk of glitch disruption and so it is these which we will focus on in this chapter. Continuous GW searches ideally employ a *fully-coherent* search which consists of matched filtering the template against the data over a coherence time T_{coh} . However, such a search is computationally intensive and so GW searches typically revert to using a *semi-coherent* method in which the total observation time T_{span} is divided into N_{seg} segments each of duration T_{coh} . Each of

these is then fully-coherently searched and then recombined to give a semi-coherent measurement which is insensitive to phase jumps between segments. This method provides more sensitive searches at fixed computing cost [142]. Typically, a semi-coherent search is performed first, then interesting candidates are *followed-up*, reducing the number of segments, aiming to finally detect the signal with a fully-coherent search; see Shaltev and Prix [153] for a two-stage follow-up procedure.

A variety of targeted and all-sky searches have already been completed, with more ongoing and planned for the advanced detector era. In Table 7.1 we list the search parameters used in some recent searches.

	f_s [Hz]	\dot{f}_s [nHz/s]	T_{coh} [hrs]	T_{span} [days]	Ref
S5 E@H all-sky	50, 1190	-2, 0.1	25	694	[3]
S5 E@H galactic center	78, 496	-71, 0	11.5	302	[1]
S5 all-sky	50, 1000	-0.89, 0	0.5	365	[2]
VSR low-freq. all-sky	20, 128	-10, 0.15	2.3	185	[7]
S5 supernova remnant (Cas A)	91, 573	-60.5, -1.6	N/A	8.4	[6]

Table 7.1: Summary of some recent all-sky and directed searches. All numbers are approximate, see the references for exact ranges. Note that S5 refers to data from the fifth LIGO science run, similarly VSR refers to data from the second and fourth Virgo science runs. For the S5 supernova remnant search we present parameters for the Cassiopeia A search only and note that this was a fully-coherent search with no semi-coherent stage.

To make practical use of the results obtained in this chapter, we will make predictions for the glitch rate and magnitude in the context of typical GW searches. As an illustrative example, we will use the search parameters that were used in LIGO’s Einstein@Home all-sky search for periodic gravitational waves in the fifth science run (S5) data [3], hereafter referred to as the E@H search. This search used the semi-coherent *Hough-transform* method [99] and is typical in terms of the timing parameters, although in future searches it is expected that the number of segments will be decreased due to improvements in computing power. We find in Section 7.4 that the effect of a glitch is independent of the sign of the frequency derivative, and so for convenience we truncate the GW spin-down parameter space to $\dot{f}_s = [-2, 0] \times 10^{-9} \text{ Hz s}^{-1}$.

7.3 Statistical properties of the observed glitch database

In this section, we will discuss the properties of glitches in the observed radio pulsar population using the glitch catalogue maintained by Espinoza et al. [59] and available at www.jb.man.ac.uk/pulsar/glitches.html. Our goal is to make a statement about how often glitches occur and their magnitudes for the types of neutron star which may be emitting GWs in the parameter space of typical GW searches. This task is made difficult since many searches look for young, rapidly rotating stars for which we only have a small sample of observations. Therefore, we must moderately extrapolate to apply our results to the observed radio pulsar population.

Radio pulsar timing methods detect glitches by fitting a piecewise Taylor expansion in the phase on either side of the event, with a modelled jump in between (see Edwards et al. [57] or Section 2.1 for a detailed discussion). The glitch catalogue reports 472 events from 165 pulsars (as of the 27th of June 2016); for each of these events a value is reported for the frequency jump $\delta\nu$ and frequency derivative $\delta\dot{\nu}$, if it can be measured. We cross-reference the glitch catalogue with the ATNF Manchester et al. [117] pulsar catalogue available at www.atnf.csiro.au/people/pulsar/psrcat/ in order to get the glitching pulsar’s timing properties.

Of the 472 listed glitches, we find 15 with no ATNF cross-reference, 1 with $\delta\nu < 0$, and 4 with no measured $\dot{\nu}$ in the ATNF catalogue; these pulsars are removed from our data set. Additionally, we find 54 glitches which have either no measured $\delta\dot{\nu}$, or a measured value consistent with 0. These will be included where possible.

7.3.1 Glitch magnitudes

Espinoza et al. [59] argued that the glitch catalogue contains glitches from two distinct sub-populations of pulsars. There is the main population with $\delta\nu$ magnitudes ranging from 10^{-9} to 10^{-5} Hz, and a second, less numerous population with larger magnitudes of $\delta\nu$. The smaller population are described as ‘Vela-like’ in that the pulsars that undergo these glitches have similar characteristic ages and magnetic fields to the Vela pulsar (B0833-45). We reproduce the evidence for this finding in Figure 7.1 where we plot the histogram of all observed $\delta\nu$ values. This demonstrates the bimodality found by Espinoza et al. [59].

To check that the bimodality is not an artefact of the histogram bin sizes we estimate the probability density function using a Gaussian kernel density estimate (KDE). Specifically we use the Jones et al. [92] implementation. This is also plotted in Figure 7.1 and demonstrates two distinct peaks, although the smaller peak could be interpreted as two modes close together.

By eye, it is clear that there are at least two modes to the histogram. However, it is also possible that there may be more modes with less distinct differences in their mean. We investigate this issue in Appendix 7.A by applying a Bayesian model comparison to Gaussian mixture models (see Chapter 18 of Gelman et al. [64] for an introduction to mixture models) varying the number of components and also allowing for a skew as described in O’Hagan and Leonard [129]. We find that all models with 2 or more components fit the data decisively better than a single component. Marginal gains are found by allowing the models to be skewed and have 4 or more components, but no single model is outstanding amongst the others. For this reason we choose to use a 2-component model with skew; this provides a good empirical description of the data and is pragmatic in that we limit the number of components to 2 for interpretability. We note that this description is empirical and we do not intend to make any substantive claim relating the two components to the two mechanisms discussed in Chapter 2.1.

Having obtained a fit to the magnitude of the glitch in frequency using a 2-component skewed-Gaussian mixture model, we use this fit to label each data point as originating

from one of the 2 skewed-Gaussian components . Specifically, to each data point we assign the label based on the maximum probabilities of each of the two components, given the maximum posterior model parameters derived in the fitting process.

The resulting mixture components and individual distributions are plotted in Figure 7.1 and in Table 7.2 we provide the resulting mean, standard deviation, weights and skewness of the two components in log-space. This method identifies the two subpopulations in a manner consistent with the observations by Espinoza et al. [59] and notably the Vela-like component suffers a significant skew.

	Mean	Std. Dev.	Weight	Skew
Normal	-8.413	1.590	0.701	1.084
Vela-like	-4.407	0.534	0.299	-9.660

Table 7.2: Log-normal mean, standard deviation and weight for the components of the mixture model as fitted in Figure 7.1.

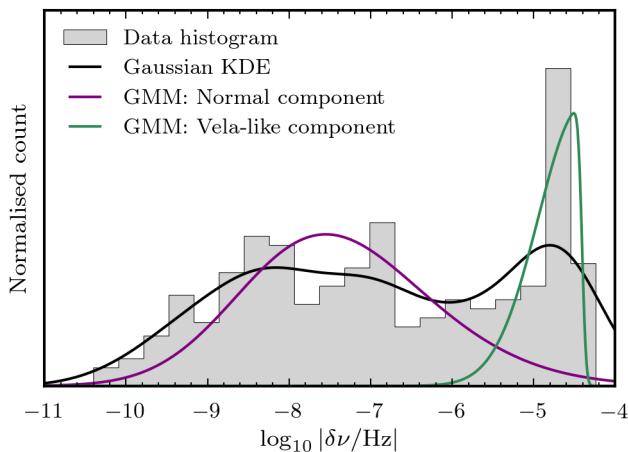


Figure 7.1: The distribution of glitch magnitudes $\delta\nu$ observed in the glitch catalogue. This is given as both a binned histogram and a Gaussian KDE, as discussed in the text. The coloured lines mark the two components of the skewed Gaussian mixture model fitted to model the bimodality.

In Figure 7.2, we plot histograms for $\delta\nu$ and $\delta\dot{\nu}$ along with the raw data in a scatter plot. We have separated the data into the individual sub-populations, as labelled by the Gaussian mixture model, and colour coded to aid the eye. Several pulsars of interest are picked out using coloured halos. It is of general interest that not all the Vela glitches are categorised by our method as Vela-like; this can be seen by looking at the distribution of Vela glitches in Figure 7.2.

7.3.2 Overview of the population of glitches

To give an overview of all observed glitches in the context of the whole population of observed radio pulsars listed in the ATNF catalogue, in Figure 7.3 we plot two copies of the familiar ν - $\dot{\nu}$ diagram. In panel A, for each pulsar which has been observed to glitch

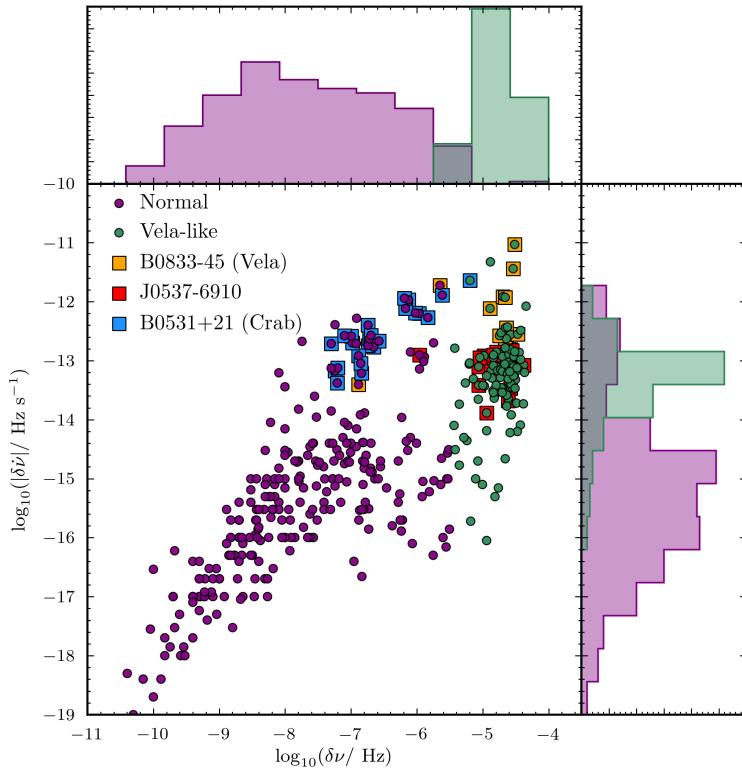


Figure 7.2: Glitch magnitudes as provided by the glitch-database [59]. This shows a scatter plot of all pairs of $\delta\nu$ and $\delta\nu$ where the colouring depends on the labelling given by the mixture model. Purple circles are the points categorised as ‘normal glitches’, while green circles are the points from the ‘Vela-like’ population. Histograms for both glitch magnitudes are also given for each sub-population. Coloured halos highlight glitches from interesting pulsars.

we add a coloured circle. Those pulsars which glitched multiple times are marked by a larger coloured circle with the area proportional to the number of glitches. In panel B, we again mark each pulsar that has been observed to glitch with a coloured circle, but here the area of the coloured halo marks the pulsar’s average glitch magnitude. For both plots, different colours have been used to partition the Vela-like and normal glitches (note that some pulsars display glitches from both populations) and a shaded box marks the parameter space searched by the S5 E@H search. Finally, dashed lines mark isolines of constant characteristic age as defined by $\tau_{\text{age}} = |\nu/\dot{\nu}|$.

While the bulk of observed glitching pulsars are from the main pulsar population, the fraction of young pulsars ($\tau_{\text{age}} < 10^5$ yrs) which glitch is proportionally higher than in the normal population. Vela-like glitches occur predominantly in the young pulsars with none seen in pulsars with $\tau_{\text{age}} > 10^7$ yrs. It is also noticeable that younger pulsars display a greater number of glitches. Note that, since we have not observed all pulsars for the same duration, one cannot infer the relative glitch rate from the number of glitches alone.

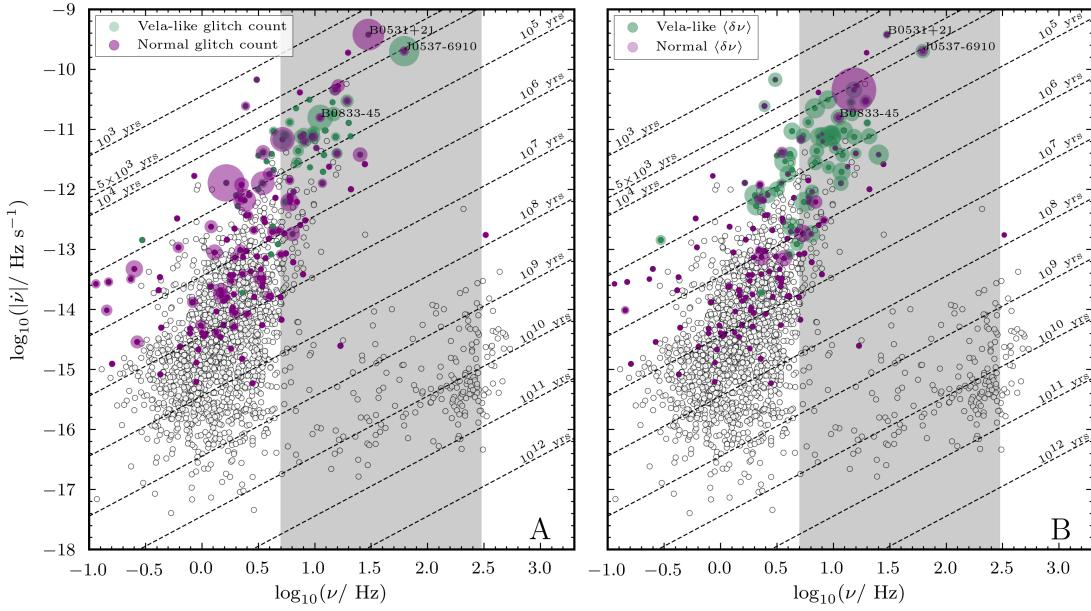


Figure 7.3: Frequency-frequency derivative plot of all pulsars in the ATNF catalogue [117]. **A:** A single coloured point marks pulsars which have been observed to glitch; the area of the coloured halo is proportional to the number of observed glitches from that pulsar. **B:** A single coloured point marks pulsars which have been observed to glitch, the area of the coloured halo is proportional to the average glitch magnitude from that pulsar. We have used purple for ‘normal’ glitches and green for ‘Vela-like’ glitches, as defined by the skewed Gaussian mixture model in Section 7.3.1. Note that, for the glitch magnitudes, the relative scaling for the Vela-like and normal populations are *not* the same since the Vela-like pulsars are significantly larger: the area representing the normal glitch magnitudes are scaled 3 times larger than the Vela-like glitch magnitudes. The grey shaded box marks the parameter space of typical GW all-sky searches which cover a rotational frequency ν range of $10 - 600$ Hz (assuming they search for signals with $f_s = 2f_s$).

For the normal-glitch population, Espinoza et al. [59] noted that pulsars with $\tau_{\text{age}} < 5 \times 10^3$ yr undergo small or medium sized glitches ($\delta\nu < 10^5$ Hz). It is postulated that the higher temperatures in younger pulsars prevents the glitch mechanism working effectively. This effect is consistent with Figure 7.3B: the pulsars with the largest average glitch sizes have $\tau_{\text{age}} \sim 10^5$ yrs, while younger pulsars tend to exhibit smaller glitches on average.

7.3.3 Extrapolating: glitch magnitudes

We would like to be able to predict the glitch magnitude for the unobserved pulsar population targeted by GW searches. In particular, we need to extrapolate up to the large values of $\dot{\nu}$ searched for in many all-sky GW searches, where few observed radio pulsars exist.

It has previously been found by McKenna and Lyne [120], Lyne et al. [114], Wang et al. [163], and Espinoza et al. [59] that the glitch activity (defined in the first of these references) correlates well with $|\dot{\nu}|$ and the characteristic age τ_{age} . We choose not to combine the rate and magnitude information together into the activity, but estimate both separately as these are of most direct relevance to GW searches.

We investigated correlations of the glitch magnitudes $\delta\nu$ and $\delta\dot{\nu}$ with the frequency, frequency-derivative and characteristic age. In Table 7.3 we present the Pearson correlation coefficient for each glitch magnitude ($\delta\nu$ and $\delta\dot{\nu}$). This is done for three groups: all the data together, then individually for the normal population and the Vela-like population. For the normal population, both glitch magnitudes most strongly correlates with

		$\log_{10} \tau_{\text{age}} $	$\log_{10} \nu $	$\log_{10} \dot{\nu} $
All	$\log_{10} \delta\nu $	-0.634	0.538	0.68
	$\log_{10} \delta\dot{\nu} $	-0.846	0.672	0.88
Normal	$\log_{10} \delta\nu $	-0.631	0.390	0.64
	$\log_{10} \delta\dot{\nu} $	-0.864	0.604	0.88
Vela-like	$\log_{10} \delta\nu $	0.035	0.12	0.040
	$\log_{10} \delta\dot{\nu} $	-0.62	0.376	0.593

Table 7.3: The correlation coefficient between the glitch magnitudes and the measured and inferred timing properties of the source pulsar. For each row, the largest value is highlighted in bold.

spin-down rate $\dot{\nu}$, although we recognise that τ_{age} is only marginally worse. In contrast, $\delta\nu$ for the Vela-like population has a weak correlation with all predictor variables, but $\delta\dot{\nu}$ does correlate well showing the strongest correlation with the characteristic age. We choose to use $\dot{\nu}$ as a predictor variable for both the normal and Vela-like populations. For the latter the correlation coefficient suggests that τ_{age} may be a better predictor, however $\dot{\nu}$ is only marginally worse and it makes it simpler to interpret later results if the same predictor is used for both populations. In practise, our results will be robust to either choice of the predictor variable.

In Figure 7.4A and Figure 7.4B we scatter-plot the glitch magnitudes against the spin-down rate of the pulsar to demonstrate the correlation. For both plots we have added coloured halos to label several interesting pulsars. These help to show that there can be almost as much variation in the glitch magnitude of a single pulsar as from the entire population.

Fitting a linear function in log-log space (see Appendix 7.B for details) our resulting fitting formulae for the frequency jump due to each separate population is

$$\begin{aligned}\langle \delta\nu \rangle_{\text{Normal}} &= 10^{-0.89} |\dot{\nu}|^{0.55} 10^{\pm 0.93}, \\ \langle \delta\nu \rangle_{\text{Vela-like}} &= 10^{-4.62} |\dot{\nu}|^{0.01} 10^{\pm 0.28}.\end{aligned}\tag{7.1}$$

and for the frequency-derivative jumps is

$$\begin{aligned}\langle \delta\dot{\nu} \rangle_{\text{Normal}} &= 10^{-4.16} |\dot{\nu}|^{0.90} 10^{\pm 0.67}, \\ \langle \delta\dot{\nu} \rangle_{\text{Vela-like}} &= 10^{-7.06} |\dot{\nu}|^{0.57} 10^{\pm 0.66}.\end{aligned}\tag{7.2}$$

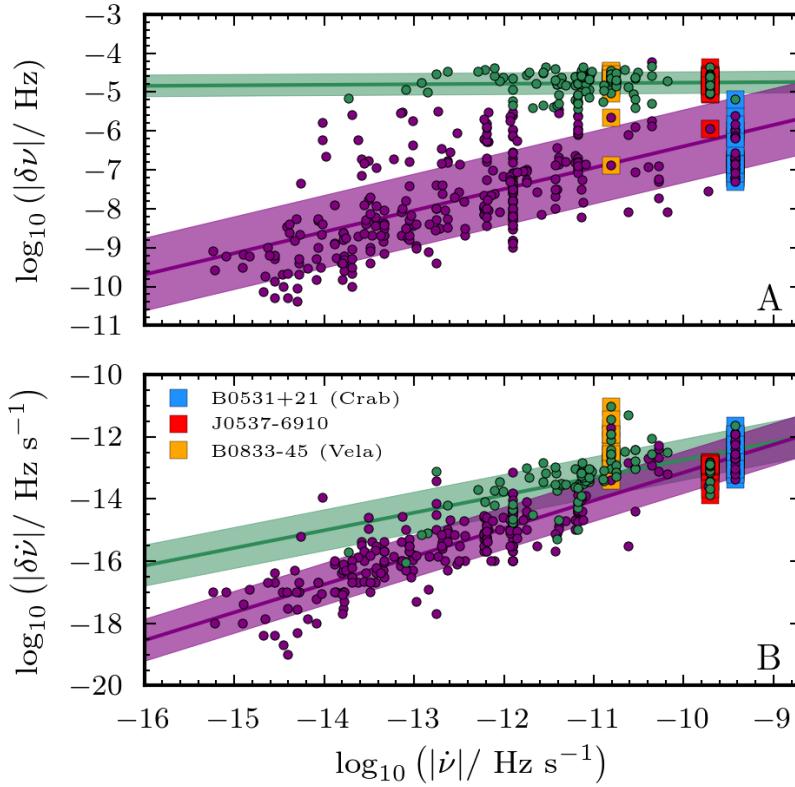


Figure 7.4: **A:** the magnitude of $\delta\nu$ as a function of the source pulsar’s spin-down rate. **B:** the absolute value of the magnitude of $\delta\nu$ against the measured spin-down rate. The coloured lines and shaded bands are the best fits from Eqn. (7.1) for $\delta\nu$, and Eqn. (7.2) for $\delta\nu$; the green lines mark the Vela-like fit while the purple lines park the fit to the normal population. Vertical clustering in the observed data points is the result of multiple glitches observed from a single source. Coloured halos highlight glitches from some interesting pulsars.

Note that the last factor here provides an estimate of the variability about the linear fit, while neglecting this term gives the mean. We plot both of these fitting formulae in Figure 7.4; the estimate of the variability is indicated by a shaded band.

Taking these fitting formulae, in Figure 7.5 we plot the predicted glitch magnitudes (in frequency and spin-down rate respectively) over the S5 E@H search parameters. We have similarly transformed the variation estimate and plotted it as a shaded band. These fits do not provide a precise statement about the magnitude of glitches, but are sufficient to estimate the order-of-magnitude that we might expect.

7.3.4 Extrapolating: average glitch rate

In order to estimate the average rate of glitches, Espinoza et al. [59] grouped pulsars by their spin-down rate $\dot{\nu}$, including pulsars which have not yet been observed to glitch. From this grouping, the authors used the measured number of glitches N_g to calculate

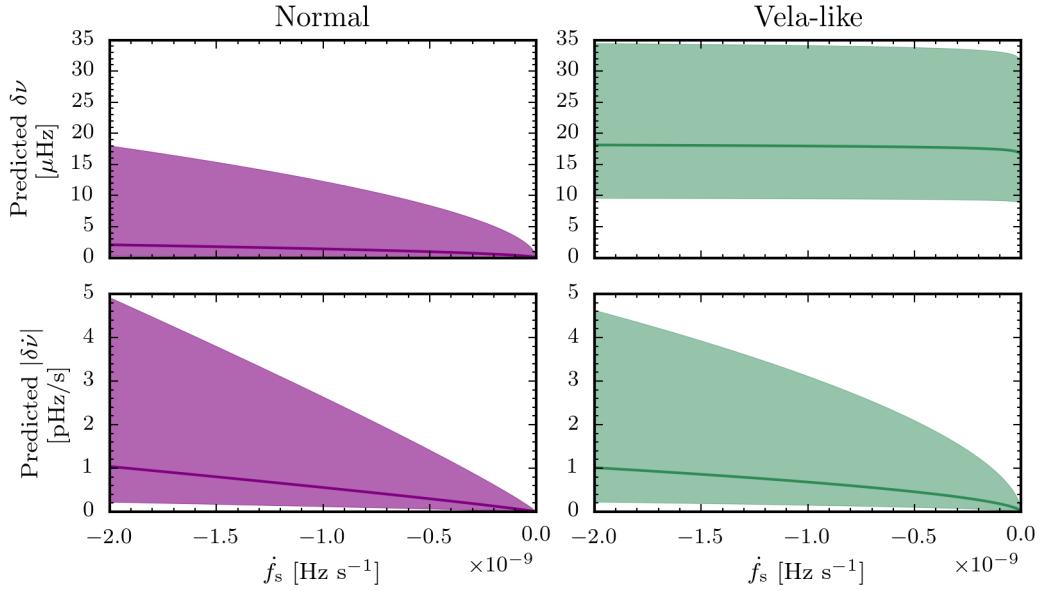


Figure 7.5: Predicted glitch magnitude of $\delta\nu$, in the upper panels, and $\delta\dot{\nu}$, in the lower panels. These are plotted over the E@H parameter space using the fitting formulae of Eqn. (7.1) and Eqn. (7.2). The shaded bands indicate the estimated error neglecting correlations.

a mean glitch rate $\langle \dot{N}_g \rangle$. In Figure 10 of their work they show that, to a good approximation, in log-space the mean glitch rate depends linearly on the spin-down rate; we reproduce this in Figure 7.6 using the data from Table 4 of Espinoza et al. [59].

In order to extrapolate, we perform a linear regression to the data in log-space following the method described in Appendix 7.B. This gives a prediction of the glitch rate as

$$\langle \dot{N}_g \rangle = 10^{-3.00} |\dot{\nu}|^{0.47} 10^{\pm 0.31} \text{ s}^{-1}, \quad (7.3)$$

where $\dot{\nu}$ is measured in Hz/s. The exponent agrees with that found by the original authors (they do not provide the prefactor).

We plot this average glitch rate in Figure 7.7A over the range of \dot{f}_s values considered in the S5 E@H search. We also multiply this rate by the span of a typical search, $T_{\text{span}} \approx 1$ yr, to obtain $\lambda(T_{\text{span}}=1\text{yr})$, the expected number of glitches during the search; this is also plotted in Figure 7.7A, on the right-hand axis.

Interpreting this average glitch rate requires a substantive physical model. Melatos et al. [123] demonstrated that glitch waiting times are consistent with an avalanche process transferring angular momentum from the core superfluid to the crust. Choosing 9 pulsars which had glitched 5 times or more, they found that 7 of these were consistent with a constant rate Poisson process such that each glitch event is statistically independent. In the remaining two, J0537-6910 and B0833-45 (Vela), they find that a quasiperiodic component coexists with the Poisson process and accounts for about 20% of the events.

Of the glitch catalogue, J0537-6910 accounts for 23 and B0833-45 (Vela) for 17 of the total 472 events. Assuming that 20% of these are due to the quasiperiodic component, this

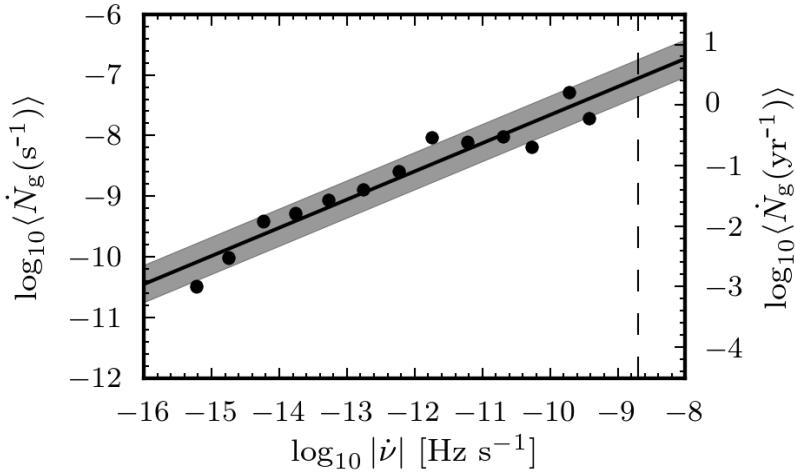


Figure 7.6: Reproduction of Figure 10 from Espinoza et al. [59] given as both log of the glitch rate per second (left axis) and per year (right axis). Black dots are the original data points, the solid line and shaded region are our best-fit line and a measure of the variability as given in Eqn. (7.3). A vertical dashed line marks -2×10^{-9} , the largest absolute spin-down rate used in the S5 E@H search.

is $\sim 1.7\%$ of the total number of observed glitches. It is possible that other pulsars also exhibit a quasiperiodic component, so the total fraction of glitches from a quasiperiodic component should be greater than 1.7% . However, we may still claim that the majority of glitches in the catalogue are due to a Poisson like process.

Assuming that all glitches used to estimate $\langle \dot{N}_g \rangle$ are due to a Poisson process, we can calculate the probability of one or more glitches occurring during a typical search lasting 1 yr as a function of \dot{f}_s . To do this, we take the estimated number of glitches during a typical search λ , as given on the right axis of Figure 7.7A, and sum the Poisson probability mass function from 1 to infinity

$$P(N_g \geq 1; T_{\text{span}}) = \sum_{N_g=1}^{\infty} \frac{\lambda^{N_g} e^{-\lambda}}{N_g!}, \quad (7.4)$$

where the dependence on T_{span} is in $\lambda = \lambda(T_{\text{span}})$. Note that in practise we truncate the summation at a finite level where the mass function is negligible.

In Figure 7.7B we plot this probability using λ from the upper plot. This suggests that for almost all of the E@H spin-down rate parameter space, it is more probable to have at least one glitch per year than to have none. This estimation will suffer bias from the inclusion of the quasiperiodic components into the calculation of $\langle \dot{N}_g \rangle$. Nevertheless, we hope it provides some quantitative measure of the probability of a glitch occurring.

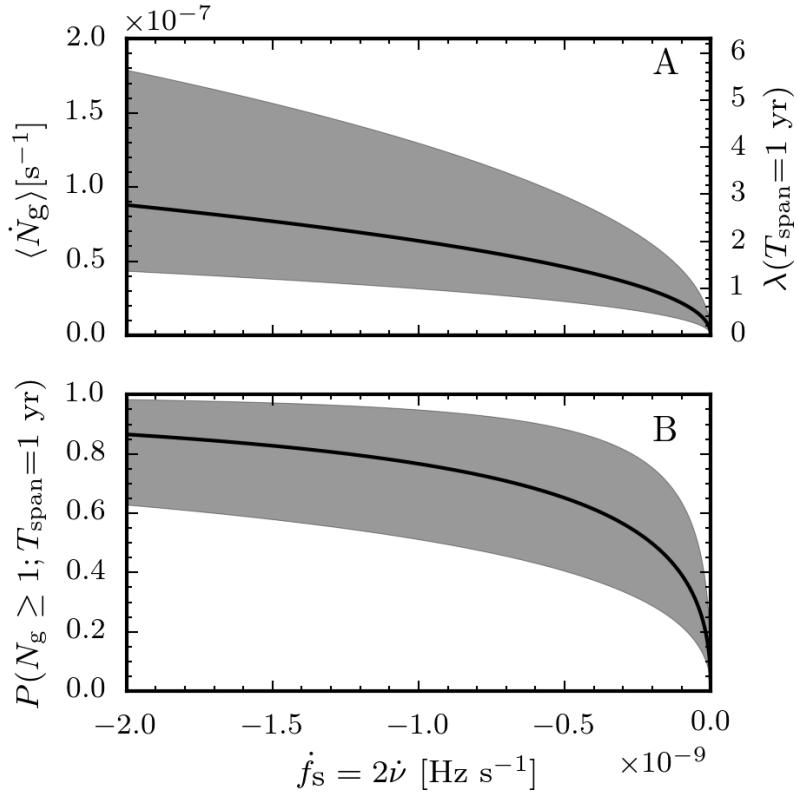


Figure 7.7: **A:** the average glitch rate against the spin-down rate \dot{f}_s over the range of values used in the S5 E@H search. On the right axis is the corresponding average number of glitches for a search of duration 1 yr. **B:** the probability of observing one or more glitches during a search lasting for 1 yr assuming a Poisson distribution.

7.4 Calculating the mismatch due to a single glitch

In Section 7.3 we established the statistical properties of observed radio pulsar glitches and provided fitting formulae for the magnitude and rate of glitches. In this section, we will convert these results into a quantified effect on fully-coherent and semi-coherent GW searches using the tools developed in Section 6.2 and Section 6.4.

For this work, we consider a glitch to consist of an instantaneous jump in the phase, frequency, and frequency derivative of the GW signal; the magnitudes of these quantities we will denote by $\delta\phi$, δf , and $\delta\dot{f}$. Since we don't yet understand the glitch mechanism and what happens during the glitch, it's unclear how meaningful it is to discuss a jump in the GW phase. Nevertheless, modelling the signal with a piecewise Taylor expansion naturally includes such a phase jump, therefore we will include it in our discussion. The size of the jump in the GW frequency and frequency derivative can be estimated from the observed jumps in Section 7.3 (assuming $\delta\dot{f} = 2\delta\nu$ etc.). Aside from these jumps there may also be an exponential relaxation in the post-glitch dynamics and a rise-time during which the glitch occurs; for now all such phenomena will be ignored, but we return to this in Section 7.5.4.

Our glitch signal can be modelled by a piecewise Taylor-expansion with two subdomains. We denote the time of the glitch as RT_{span} where R is a dimensionless fractional quantity such that $R \in [0, 1]$ and T_{span} is the observation time of the search. Then labelling the period before the glitch as A and after as B , the parameter offsets may be written

$$\Delta\lambda^{\alpha a}(t) = \begin{cases} \Delta\lambda^{\alpha A} & \text{if } t < RT_{\text{coh}} \\ \Delta\lambda^{\alpha B} & \text{if } t > RT_{\text{coh}} \end{cases}, \quad (7.5)$$

where $\Delta\lambda^{\alpha j} = \lambda_s^{\alpha j} - \lambda_t^{\alpha j}$. We will set the reference times for each subdomain half-way through the subdomain and also define a global reference time t_{ref} at the glitch, RT_{coh} . To help orient the reader with these choices, we provide a schematic of the signal frequency over the glitch in Figure 7.8.

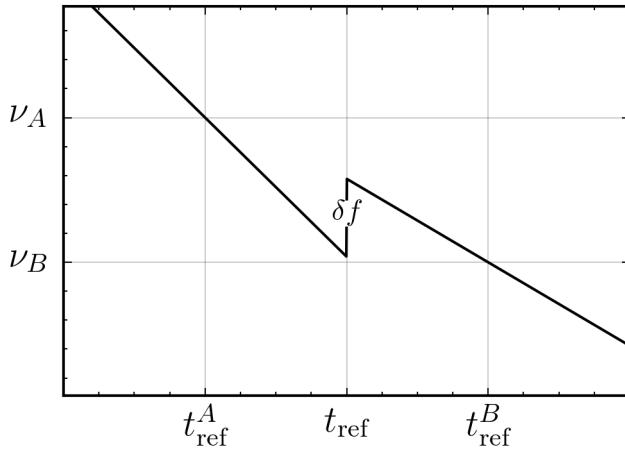


Figure 7.8: Illustration of the references times and frequency jump over the glitch. Note that the global reference time t_{ref} is set to coincide with the time at which the glitch occurs; in this instant it is half-way through, but could in general be at any time during the observation. The local reference times are set halfway through each subdomain.

The jumps at the glitch itself may then be parameterised as

$$\begin{aligned} \delta\phi &= \phi_s^B(RT_{\text{coh}}) - \phi_s^A(RT_{\text{coh}}) \\ \delta f &= f_s^B(RT_{\text{coh}}) - f_s^A(RT_{\text{coh}}) \\ \dot{\delta f} &= \dot{f}_s^B - \dot{f}_s^A \end{aligned} \quad (7.6)$$

The time-dependence here indicates that one must account for the changes due to lower order terms between the reference time, halfway through the subdomain, and the glitch. Since we do not consider second-order spin-down terms, this is not required for the $\delta\dot{f}$ calculation.

In the following sections, we consider fully-coherent and semi-coherent GW searches for this signal which contains a single glitch. In general these searches are performed over a grid of Taylor expansion templates in the frequency and spin-down parameter space and then, if some subset of the templates return a sufficiently high SNR, the template with

the smallest mismatch is taken as a candidate. This method minimises the mismatch with respect to the search parameters f_t and \dot{f}_t ; we will include this minimisation step in the calculation.

7.4.1 A single glitch in a fully-coherent search

7.4.1.1 A fully-coherent search: analytic metric-mismatch

To calculate the fully-coherent mismatch, we first expand the summation in Eqn. (6.49) over the two subdomains giving three terms

$$\tilde{\mu} = g_{\alpha\beta AA} \Delta\lambda^{\alpha A} \Delta\lambda^{\beta A} + 2g_{\alpha\beta AB} \Delta\lambda^{\alpha A} \Delta\lambda^{\beta B} + g_{\alpha\beta BB} \Delta\lambda^{\alpha B} \Delta\lambda^{\beta B}. \quad (7.7)$$

To calculate the metric components, we use Eqn. (6.56) with $\Delta T^A = RT_{\text{coh}}$ and $\Delta T^B = (1 - R)T_{\text{coh}}$. We then insert the parameter offsets $\Delta\lambda^{\alpha a}$ from Eqn. (7.5), and explicitly write the mismatch in terms of the components of $\Delta\lambda^{\alpha a} = [\Delta\phi^a, \Delta f^a, \Delta \dot{f}^a]$. After some manipulation, the mismatch between the glitch signal and an arbitrary Taylor expansion template is

$$\begin{aligned} \tilde{\mu} = & R(1 - R)(\Delta\phi^A - \Delta\phi^B)^2 + \frac{\pi^2 T_{\text{coh}}^2}{3} [R^3(\Delta f^A)^2 + (1 - R)^3(\Delta f^B)^2] \\ & + \frac{\pi^2 T_{\text{coh}}^4}{720} [R^5(9 - 5R)(\Delta \dot{f}^A)^2 - 10R^3(1 - R)^3\Delta \dot{f}^A \Delta \dot{f}^B + (1 - R)^5(5R + 4)(\Delta \dot{f}^B)^2] \\ & + \frac{\pi T_{\text{coh}}^2}{6} [R(1 - R)(\Delta\phi^B - \Delta\phi^A)((1 - R)^2\Delta \dot{f}^B - R^2\Delta \dot{f}^A)]. \end{aligned} \quad (7.8)$$

The first three terms are the independent contributions to the mismatch from the jumps in phase, frequency, and spin-down; the last term is the mixture term between the phase and the spin-down. This expression has a maximum when $R = 1/2$, and collapses to the usual mismatch between two smooth Taylor expansions when R is 0 and 1. Furthermore, we verified that when the magnitude of the glitch is zero, one recovers the usual fully-coherent mismatch between two smooth Taylor expansions.

The mismatch in Eqn. (7.8) is a function of the individual signal parameter $\lambda_s^{\alpha a}$, and the template parameters $\lambda_t^{\alpha a}$. We are not interested in arbitrary choices of the template parameters, but those which minimise the mismatch as would be found by searching over a small area in parameter space and selecting the template with the smallest mismatch as a detection candidate. We therefore analytically minimise Eqn. (7.8) with respect to the template parameters f_t and \dot{f}_t . We find that

$$\begin{aligned} f_t^{\min} = & f_s^A + \delta f(1 - R)^2(2R + 1) + \frac{3\delta\phi}{\pi T_{\text{coh}}} R(1 - R) \\ & + \frac{T_{\text{coh}}}{2} (\dot{f}_s^A(1 - R) + \delta\dot{f}(1 - R)^3(1 + R)), \end{aligned} \quad (7.9)$$

and

$$\begin{aligned}\dot{f}_t^{\min} = & \dot{f}_s^A + \delta\dot{f}(1-R)^3(6R^2+3R+1) \\ & + \frac{30\delta f}{T_{\text{coh}}}R^2(1-R)^2 + \frac{30\delta\phi}{\pi T_{\text{coh}}^2}R(1-R)(2R-1),\end{aligned}\quad (7.10)$$

which are expressed at the global reference time t_{ref} .

Inserting these into Eqn. (7.8) and simplifying yields a minimum fully-coherent mismatch of

$$\begin{aligned}\tilde{\mu} = & R(1-R)(4R(1-R)(5R(1-R)-2)+1)\delta\phi^2 \\ & + \frac{\pi^2 T_{\text{coh}}^2}{3}R^3(1-R)^3(4-15R(1-R))\delta f^2 \\ & + \frac{\pi^2 T_{\text{coh}}^4}{5}R^5(1-R)^5\delta\dot{f}^2 \\ & + 2\pi T_{\text{coh}}R^2(1-R)^2(1-2R)(5R(1-R)-1)\delta f\delta\phi \\ & + \pi^2 T_{\text{coh}}^3 R^4(1-R)^4(2R-1)\delta\dot{f}\delta f \\ & + \frac{\pi T_{\text{coh}}^2}{3}R^3(1-R)^3(2-12R(1-R))\delta\phi\delta\dot{f}.\end{aligned}\quad (7.11)$$

An important distinction must be made here between Δf^a which is the frequency offset between the signal and template in the a^{th} subdomain, and δf the signal frequency jump at the glitch. Notably, the minimum mismatch depends only on $\delta\phi$, δf , and $\delta\dot{f}$; it is independent of the overall phase, frequency, or spin-down of the signal.

Since the probability distribution of R should be uniform over the search duration, we can average Eqn. (7.11) over R to get the expectation

$$\langle\tilde{\mu}\rangle_R = \frac{3}{70}\delta\phi^2 + \frac{T_{\text{coh}}^2}{630}\left(\pi^2\delta f^2 - \pi\delta\phi\delta\dot{f}\right) + \frac{\pi^2 T_{\text{coh}}^4}{13860}\delta\dot{f}^2. \quad (7.12)$$

The mixture terms $\delta f\delta\phi$ and $\delta f\delta\dot{f}$ vanish in the averaging process.

7.4.1.2 Simple estimates

We now make some rough estimates based on Eqn. (7.12), the R -averaged mismatch for a fully-coherent search. Firstly, let us consider a glitch which consists of a jump solely in the phase $\delta\phi$. For such a glitch, we can calculate the size of a phase-jump which would produce a mismatch of $\tilde{\mu} = 0.1$

$$\delta\phi = \sqrt{\frac{70\tilde{\mu}}{3}} \approx 1.5 \text{ rad} \left(\frac{\tilde{\mu}}{0.1} \right) \quad (7.13)$$

For the frequency and spin-down rate jumps, a simple way to quantify the significance of a given glitch is to ask ‘over what coherence time would a fully-coherent search accumulate a mismatch of $\tilde{\mu} = 0.1$?’. We will consider each type of jump independently,

since we are only hoping to make order of magnitude estimates. For the frequency jump we find that

$$T_{\text{coh}} = \frac{\sqrt{630\tilde{\mu}}}{\pi\delta f} \approx 2.9 \text{ days} \left(\frac{\tilde{\mu}}{0.1} \right)^{\frac{1}{2}} \left(\frac{\delta f}{10^{-5} \text{ Hz}} \right)^{-1}, \quad (7.14)$$

while for the spin-down rate jump

$$T_{\text{coh}} = \left(\frac{13860\tilde{\mu}}{\pi^2\delta\dot{f}^2} \right)^{\frac{1}{4}} \approx 40 \text{ days} \left(\frac{\tilde{\mu}}{0.1} \right)^{\frac{1}{4}} \left(\frac{\delta\dot{f}}{10^{-12} \text{ s}^{-2}} \right)^{-\frac{1}{2}}. \quad (7.15)$$

We have parameterised using the largest jumps seen in the glitch-catalogue as can be seen in Figure 7.2. This is a useful order-of-magnitude estimate and tells us that over timescales comparable to current and future searches (at least the fully-coherent follow-up) the mismatch can potentially rise above 0.1. We will investigate the predictions of Eqn. (7.12) in depth later on in Section 7.5.

7.4.2 A single glitch in a semi-coherent search

Having investigated the mismatch for a glitch in a fully-coherent search, we now consider the same glitch, but in a semi-coherent search as introduced in Section 6.2.5. The important point to recall is that the semi-coherent segments must be summed along the same point in parameter space. In practise, this means all values of the template parameters λ_t^α must be equal when expressed at the same reference time.

7.4.2.1 A semi-coherent search: analytic metric-mismatch

For a semi-coherent search the observation time T_{span} is divided into N_{seg} equal length segments of duration T_{coh} . We then compute the fully-coherent mismatch in each segment and then sum the squared SNR along the template parameters to get the squared semi-coherent SNR. To calculate the semi-coherent mismatch, we can make the simplifying assumption that the glitch occurs exactly at the interface between two segments such that RN_{seg} is an integer, where R measures the fraction of the observation period at which the glitch occurs. We will derive the mismatch under this assumption and then test how and when it breaks down using numerical simulations.

Under this assumption, in each segment both the signal and the template are Taylor expansions and so we can use the Brady et al. [35] formalism described in Section 6.2.4. We do not need to use the generalised metric-mismatch developed in Section 6.4.

For each segment we distinguish between the local reference time t_{ref}^j for the j^{th} segment and t_{ref} the global reference time, taken to be at the glitch RT_{span} . Then, in the j^{th} segment, the parameter offsets are calculated by transforming the global parameters to

the local offset:

$$\Delta f^j = \begin{cases} f_t - f_s^A + (\dot{f}_t - \dot{f}_s^A)(t_{\text{ref}}^j - t_{\text{ref}}) & \text{if } j \leq RN_{\text{seg}} \\ f_t - f_s^B + (\dot{f}_t - \dot{f}_s^B)(t_{\text{ref}}^j - t_{\text{ref}}) & \text{if } j \geq RN_{\text{seg}} \end{cases}, \quad (7.16)$$

$$\Delta \dot{f}^j = \begin{cases} \dot{f}_t - \dot{f}_s^A & \text{if } j \leq RN_{\text{seg}} \\ \dot{f}_t - \dot{f}_s^B & \text{if } j \geq RN_{\text{seg}} \end{cases}, \quad (7.17)$$

where $t_{\text{ref}}^j = T_{\text{coh}}(j - \frac{1}{2})$. We do not need to consider the phase jump at the glitch, since the fully-coherent mismatch is insensitive to an overall phase-offset; if the glitch did not occur at the interface between two segments then this would not be the case. The template parameters f_t and \dot{f}_t are the same in each segment; this reflects the fact that the semi-coherent search sums the SNR along a smooth Taylor expansion template.

We calculate the metric for a fully-coherent search with the reference time half-way through from Eqn. (6.12); then for the j^{th} segment, we calculate the fully-coherent mismatch to be

$$\tilde{\mu}^j = \frac{\pi^2 T_{\text{coh}}^2}{3} (\Delta f^j)^2 + \frac{\pi^2 T_{\text{coh}}^4}{180} (\Delta \dot{f}^j)^2 \quad (7.18)$$

Inserting this into Eqn. (6.17) we average over all the segments with the parameter offsets given by Eqn. (7.16) and Eqn. (7.17), to calculate the semi-coherent mismatch $\hat{\mu}$; for brevity we do not provide the result here.

Next, we minimise $\hat{\mu}$ with respect to the global f_t and \dot{f}_t to select the minimum mismatch. The minimising values are given by

$$\begin{aligned} f_t^{\min} = & f_s^B + \frac{R(5N_{\text{seg}}^2(R(9-6R)-4)+4)}{5N_{\text{seg}}^2-4} \delta f \\ & + \frac{2R(1-R)}{5N_{\text{seg}}^2-4} (5N_{\text{seg}}^2 R(1-R)-1) \delta \dot{f} T_{\text{coh}}, \end{aligned} \quad (7.19)$$

and

$$\begin{aligned} \dot{f}_t^{\min} = & \dot{f}_s^B + \frac{R(5N_{\text{seg}}^2 R(3-2R)+4)}{5N_{\text{seg}}^2-4} \delta \dot{f} \\ & + 30 \frac{R(1-R)N_{\text{seg}}}{T_{\text{coh}}(5N_{\text{seg}}^2-4)} \delta f. \end{aligned} \quad (7.20)$$

Inserting these back into the mismatch $\hat{\mu}$, we calculate the minimised mismatch and express it in terms of the glitch parameters δf and $\dot{\delta f}$

$$\begin{aligned}\hat{\mu} = & R(R-1) \frac{5N_{\text{seg}}^2(3R(1-R)-1)+4}{15N_{\text{seg}}^2-12} \pi^2 T_{\text{coh}}^2 \delta f^2 \\ & + R(1-R)\pi^2 T_{\text{coh}}^4 \dot{\delta f}^2 \\ & \times \frac{25N_{\text{seg}}^4 R^2(1-R)^2+N_{\text{seg}}^2(5R(R-1)-5)+4}{225N_{\text{seg}}^2-180} \\ & + \frac{5N_{\text{seg}}^3 R^2(1-R)^2}{15N_{\text{seg}}^2-12}(2R-1)\pi^2 T_{\text{coh}}^3 \delta f \dot{\delta f}.\end{aligned}\quad (7.21)$$

Assuming that the glitch should occur with the same probability at any point during the observation, we average over R and take the limit of N_{seg} being large. This gives

$$\langle \hat{\mu} \rangle_R = \frac{\pi^2 T_{\text{coh}}^2}{45} \delta f^2 + \frac{\pi^2 T_{\text{coh}}^2 T_{\text{span}}^2}{1260} \dot{\delta f}^2. \quad (7.22)$$

In the last step we substituted $N_{\text{seg}} = T_{\text{span}}/T_{\text{coh}}$.

7.4.2.2 Checking the validity of the assumption

Eqn. (7.21) relies on the assumption that the glitch occurs exactly at the interface between two semi-coherent segments. This assumption is valid in the limit for which $N_{\text{seg}} \gg 1$, but will be imprecise for small numbers of segments. To demonstrate this, in Figure 7.9 we plot the fractional absolute difference between the R -averaged metric-mismatch approximation $\langle \hat{\mu} \rangle_R$, as given in Eqn. (7.22), and the R -averaged exact mismatch $(\langle \hat{\mu} \rangle_R)_{\text{exact}}$ which we calculate numerically. To test both of the terms in Eqn. (7.22) we plot the residual difference for two glitches: one in the frequency and the other in the spin-down rate, the exact values are given in the legend. This figure shows that for $N_{\text{seg}} > 10$ the error due to our assumption that the glitch occurs exactly at the interface between two segments of the semi-coherent search decreases. The error which persists for $N_{\text{seg}} \gg 1$ is due to the metric-mismatch approximation which is expected and understood.

7.4.2.3 Simple estimates

Taking Eqn. (7.22), the R -averaged semi-coherent metric-mismatch, we can estimate the maximum coherence time for a semi-coherent search allowing for a mismatch $\mu = 0.1$. For the jumps in frequency this gives

$$T_{\text{coh}} = \frac{\sqrt{45\hat{\mu}}}{\pi \delta f} \approx 0.782 \text{ days} \left(\frac{\hat{\mu}}{0.1} \right)^{1/2} \left(\frac{\delta f}{10^{-5} \text{ Hz}} \right)^{-1}, \quad (7.23)$$

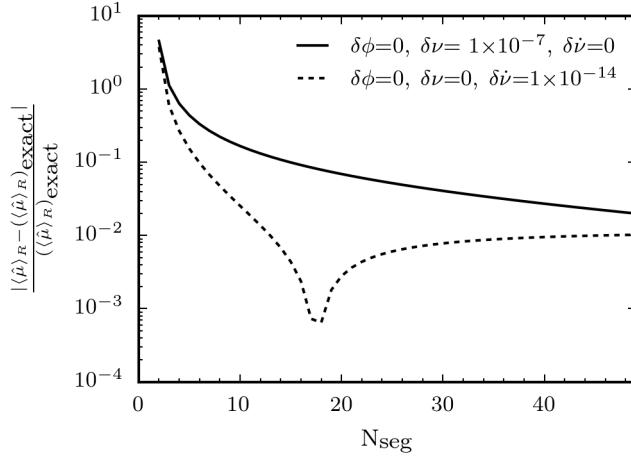


Figure 7.9: Fractional difference between the R -averaged metric-mismatch approximation $\langle \hat{\mu} \rangle_R$, as given in Eqn. (7.22), and the R -averaged exact mismatch $(\langle \hat{\mu} \rangle_R)_{\text{exact}}$. The exact mismatch is calculated numerically and does not suffer from either the metric-mismatch approximation or the assumption that the glitch occurs exactly at the interface between two segments of the semi-coherent search.

while for the jumps in frequency-derivative this gives

$$\begin{aligned} T_{\text{coh}} &= \frac{\sqrt{1260\hat{\mu}}}{\pi\delta\dot{f}T_{\text{span}}} \\ &\approx 0.689 \text{ days} \left(\frac{\hat{\mu}}{0.1} \right)^{\frac{1}{2}} \left(\frac{T_{\text{span}}}{694 \text{ days}} \right)^{-1} \left(\frac{\delta\dot{f}}{10^{-12} \text{ Hz}} \right)^{-1} \end{aligned} \quad (7.24)$$

These values are less than the coherence time used in the S5 E@H search, which was 25 hrs. These simple estimates therefore suggest that for the largest observed glitches, semi-coherent searches may suffer a non-negligible mismatch. We will investigate this and the fully-coherent mismatch case more rigorously in the next section.

7.5 Predicting the mismatch and rate of glitches in gravitational wave searches

In this section, we will use the fitting formulae of Section 7.3.1 to predict the magnitude of glitches for a range of parameters typical of all-sky searches. We will use the tools of Section 7.4 to transform the magnitudes into an estimate of the fully-coherent and semi-coherent mismatch assuming a single glitch occurred during the search. However, this mismatch does not give a complete picture of the risk, since we must also consider the predicted rate of glitches. Specifically, in converting a predicted glitch magnitude into an estimate for the fully-coherent and semi-coherent mismatch we will make use of Eqn. (7.12) and Eqn. (7.22): the R -averaged mismatch assuming that a *single* glitch occurs uniformly and at random during the observation time. We can therefore take these results as a lower bound when there is likely to be one or more glitches. On the

other hand, in regions where the chances of a glitch are low, a large mismatch only indicates that the signal would be lost in the rare event that a glitch had occurred. To provide the reader with both pieces of this puzzle, we will present results on both the expected mismatch due to a single glitch and the probability of one or more glitches occurring. This will be done for each population of glitches separately. Let us define

$$\lambda_{\text{normal}}(\dot{\nu}, T_{\text{span}}) = w_{\text{normal}} \langle \dot{N}_g \rangle T_{\text{span}} \quad (7.25)$$

$$\lambda_{\text{Vela-like}}(\dot{\nu}, T_{\text{span}}) = w_{\text{Vela-like}} \langle \dot{N}_g \rangle T_{\text{span}}, \quad (7.26)$$

as the expected number of normal and Vela-like glitches where w_{normal} and $w_{\text{Vela-like}}$ are the weights of the two populations as given in Table 7.2 and $\langle \dot{N}_g \rangle$ is the fitted glitch rate as a function of $\dot{\nu}$ which is given by Eqn. (7.3). Notice that we have made a prior specification here that the proportion of normal and Vela-like pulsars in the target population is the same as in the observed population. There is some evidence that in fact the proportion of Vela-like pulsars increases with $\dot{\nu}$; this could be modelled by a $\dot{\nu}$ -dependent weighting, however we will ignore this effect here. Eqn. (7.25) and Eqn. (7.26) can be transformed into the probability of one glitch or more occurring during the search by substitution into Eqn. (7.4). In the following discussion it is this probability which we will report on.

We note that in reality, the true mismatch for fully coherent searches will always be bounded by $[0, 1]$, but the results of this section can exceed this since we are using the metric-mismatch approximation. Similarly, for the semi-coherent searches, it is in fact bounded by the minimum of R and $1 - R$; this can be seen by realising that if the glitch is sufficiently large, the minimum mismatch is achieved by fitting to the semi-coherent segments on only one side of the glitch. Here, we report the simpler results of naively applying Eqn. (7.12) and Eqn. (7.22). The results can therefore be interpreted by realising that if the metric-mismatch is greater than 1, there will be large true mismatch, although the exact value will depend on where that glitch occurred during the observation span.

7.5.1 Fully-coherent searches

In this section, we present results on the expected mismatch and expected number of glitches for fully-coherent searches. In Figure 7.10 we plot contours of fixed mismatch as a function of the coherence time T_{coh} and the GW spin-down rate \dot{f}_{s} . Alongside are plotted the number of glitches as a function of the observation time and spin-down rate.

These figures illustrate that the predicted mismatch depends on the source population. For the normal population, larger absolute spin-down rates are associated with larger glitch magnitudes; as a result larger absolute spin-down rates are predicted to suffer more severe mismatches. For the Vela-like population, the mismatch is largely independent of $\dot{\nu}$, this is because the fitting formulae Eqn. (7.1) finds little variation in the glitch size with the spin-down rate.

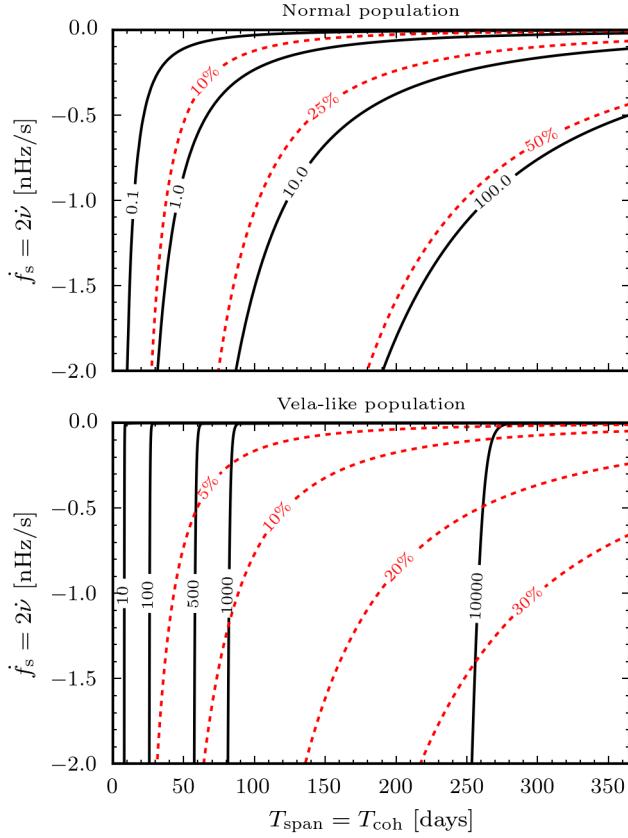


Figure 7.10: The estimated fully-coherent mismatch as a function of T_{coh} and \dot{f}_s . In the top panel is the prediction based on the fit to the normal population, in the bottom panel is the prediction based on the fit to the Vela-like population. Solid black lines provide contours of fixed mismatch. Red dashed lines have been added to indicate the probability (as a percentage) of one or more normal glitch occurring calculated by substitution of Eqn. (7.25) (for the normal population) or Eqn. (7.26) (for the Vela-like populations) into Eqn. (7.4).

For long observation times the mismatch can be severe, however this is only a concern if a glitch occurs. The expected number of glitches in both populations is, for most of the parameter space, lower than 1, but not vanishingly so.

We note here that these plots only show the ‘best-fit’ contour lines predicted by our fitting for glitch sizes. There is substantial variability in the population glitch sizes and rate, this is reflected in large error bars in the fitting formulae derived in Section 7.3. In Appendix 7.C we translate these estimates into uncertainties on the mismatch and rate contour lines.

7.5.2 Semi-coherent searches

For the semi-coherent search we repeat the process of predicting the glitch magnitudes, but then use Eqn. (7.22) to transform them into predictions for the semi-coherent mismatch. A semi-coherent search has two fixed times: the total observation span T_{span}

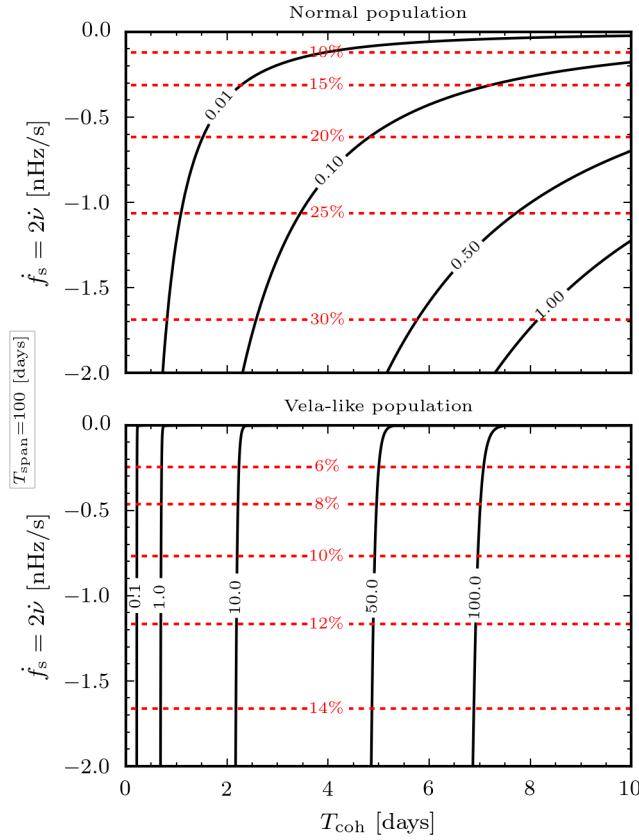


Figure 7.11: The estimated semi-coherent mismatch as a function of T_{coh} and \dot{f}_s for a fixed observation time $T_{\text{span}} = 100$ days; see Figure 7.10 for a description of the contour lines.

and the coherent segment length T_{coh} . In Figure 7.11 and Figure 7.12 we present results at fixed observation times of 100 days and 365 days and show contours of the expected semi-coherent mismatch and number of glitches as a function of the spin-down rate and coherence time.

For the semi-coherent mismatch, the contour lines tell much the same story as the fully-coherent search. The normal population produces a lower overall level of mismatch with strong \dot{f}_s dependence while the Vela-like population has larger mismatches with weak dependence on \dot{f}_s .

As in Figure 7.10, red dashed lines have been added to indicate the probability of one or more normal glitches occurring, this is calculated by substitution of Eqn. (7.25) (for the normal population) or Eqn. (7.26) (for the Vela-like population) into Eqn. (7.4). Note these are not a function of T_{coh} since we multiply the glitch rate per unit time by the fixed observation period T_{span} .

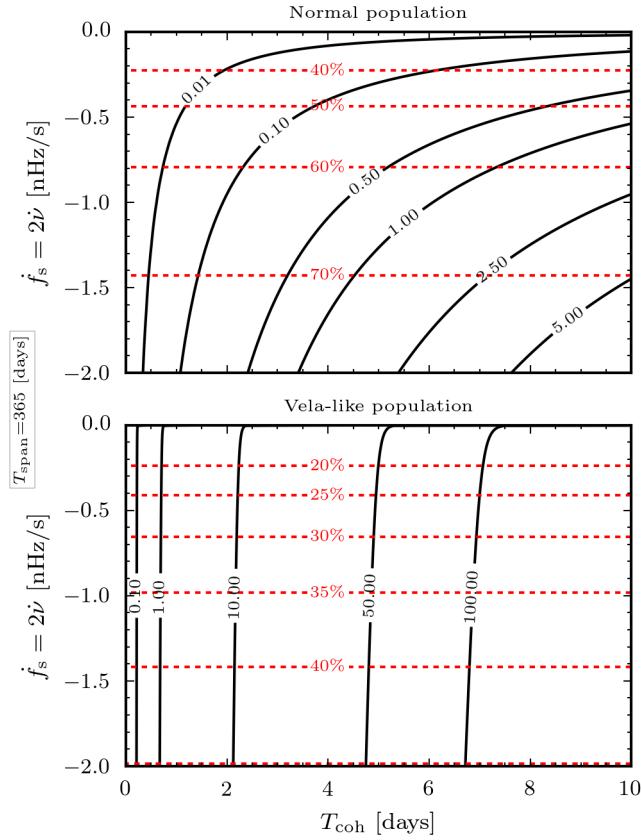


Figure 7.12: The estimated semi-coherent mismatch as a function of T_{coh} and \dot{f}_s for a fixed observation time $T_{\text{span}} = 365$ days; see Figure 7.10 for a description of the contour lines.

7.5.3 The follow-up stage

A semi-coherent search begins by searching with N_{seg} segments to find candidates. Following successful identification of such candidates, these are followed-up by semi-coherent searches with fewer segments, zooming in to constrain the parameter space. We will now show how a signal can be identified in the initial step, but subsequently lost in the follow-up, even if the glitch is quite small. For this exercise, we will investigate a signal that contains a single glitch with $\delta f = 5 \times 10^{-7}$ Hz: this is a fairly typical normal glitch size when compared to the observed normal population (see for example Figure 7.2).

In Figure 7.13 we plot the semi-coherent mismatch for this glitch as a function of the number of segments at a fixed observation time; this is inversely proportional to the coherence time which we also give on the upper axis. This is predicted analytically from the R -averaged Eqn. (7.22). We also performed a Monte-Carlo numerical simulation in which the mismatch was calculated exactly for a fixed glitch size, but an R chosen uniformly throughout the observation period. The resulting mismatches are then histogrammed and the density is shaded on to Figure 7.13: this demonstrates the spread of

mismatches about the average value. Notably there is a tight band centred on the analytic R -averaged prediction of Eqn. (7.22); there is a low-density of mismatches below this band. These occur when R is close to 0 or 1.

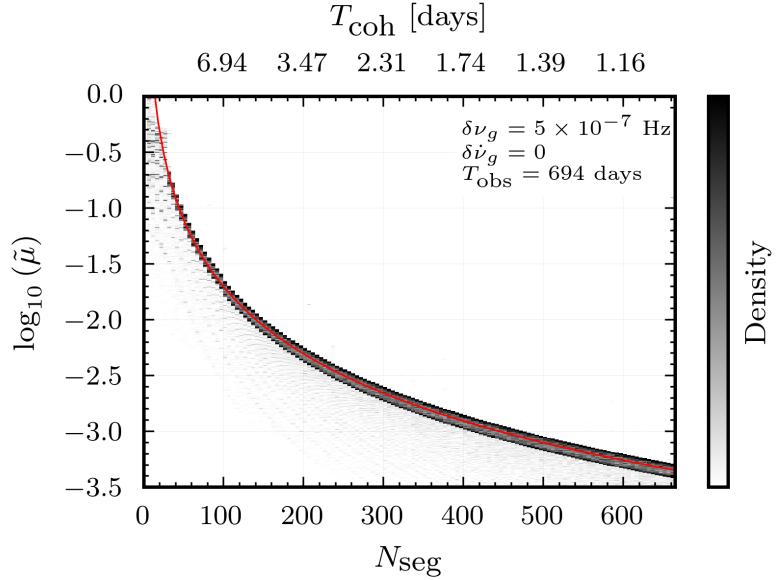


Figure 7.13: An example of how the semi-coherent mismatch changes with N_{seg} for a fixed glitch size. The red line is the prediction of the R -averaged Eqn. (7.22). The density map is constructed by performing numerical Monte-Carlo simulations over R and binning the mismatches into histograms.

This plot shows that for this glitch during the initial search at $N_{\text{seg}} = 667$, the number of segments used in the E@H all-sky search, the mismatch would be negligible $\sim 10^{-3.5}$. As a result the signal would be classified as a candidate and subsequently followed-up. We see then that, as the number of segments is decreased, the mismatch rapidly increases and hence the candidate would be dismissed.

7.5.4 Including the recovery from glitches

In this work, we have used the glitch catalogue maintained by Espinoza et al. [59] which provides $\delta\nu$, the instantaneous frequency increase at the glitch. In addition to this, some glitches also undergo a short-term exponential relaxation of some fraction of the total glitch magnitude over times of tens to hundreds of days. Lyne et al. [114] characterised the fractional recovery by $Q = \sum \Delta f_i / \delta\nu$ where $\sum \Delta f_i$ is the total glitch magnitude summed over the number of relaxation components. For each component, Δf_i is recovered over an e-folding timescale τ_i^d . In most cases, only a single component is identified.

This may have an important effect on our estimates since, if a large fraction of the glitch is recovered in a timescale short compared to the observation time, we will overestimate the mismatch. On the other hand, the exponential relaxation itself could also cause a mismatch which would tend to vanish when $\tau_i^d \ll T_{\text{span}}$, but be maximal when

$\tau_i^d \sim T_{\text{span}}$. Ideally we would like to model the exponential relaxation in detail, but the more practical approach that we will take is to assume that the relaxation time is zero, such that the actual long-term glitch magnitude is $(1 - Q)\delta\nu$; the results in this work can be considered to correspond to $Q = 0$. The effect of this is that, since $\mu \propto \delta\nu^2$, all mismatches can be rescaled by a factor $(1 - Q)^2$, given the size of Q .

To determine the appropriate size of Q that we may expect, we can use the analysis by Lyne et al. [114]. In general, it is found that for older pulsars, with lower spin-down rates, $Q \ll 1$, but Q correlates well with $\log_{10}|\dot{\nu}|$. Taking all measured Q values for the Crab pulsar [114, 164, 162], which has the highest spin-down rate -0.38 nHz/s in the analysis, we find an average and standard-deviation of 0.77 ± 0.2 for Q and 9.45 ± 6.7 days for the relaxation timescale. We note that in three cases the same glitch was measured with differing values of Q and relaxation time; this presumably reflects the fact that the measurements are uncertain and depend on the data span used. In contrast, the Vela pulsar, which has the second largest spin-down rate 0.016 nHz/s has an average Q of 0.08 for the glitches in Lyne et al. [114], with none exceeding 0.2 .

It is unclear what Q we might expect for the target population of all-sky searches, but we can make an estimate by taking $Q = 0.8$. Given this and assuming the relaxation time is much shorter than the observation period, this would mean the mismatch labels of the black contour lines in Figure 7.10, Figure 7.11, and Figure 7.12 would decrease by a factor ~ 0.04 , ultimately meaning that longer coherence times would be more robust to typical glitches. Nevertheless, even given this, many searches would still be at risk to lost signals due to the presence of a glitch. These estimates are a ‘best-case scenario’ in that they assume $\tau_i^d = 0$. A more detailed analysis is required to understand how a recovery time which is comparable to the observation time changes these estimate. However, it seems likely that it will tend to increase the level of mismatch as compared to the simple rescaling argument given here and not decrease it further.

7.5.5 Application to past and future searches

In Table 7.4 we present the expected number of glitches and the metric-mismatch (normal and Vela like) for the searches listed in Table 7.1. This is done taking the *worst-case* scenario of the largest absolute spin-down rate and hence largest glitches one might expect with no glitch recovery. Note that this is reporting the metric-mismatch which is unbounded above: a large metric-mismatch (greater than one) indicates a large mismatch.

This table gives a clear picture that for all searches in the fully-coherent follow-up, the worst mismatches are greater than one. It is true that the probability of a glitch is less than one, but not by a sufficient margin to consider them unaffected. However, the mismatch in the semi-coherent stage can be sufficiently small to be immune to glitches, provided the coherence times and observation times are small enough.

	normal population			Vela-like population		
	$\langle N \rangle$	$\langle \tilde{\mu} \rangle$	$\langle \hat{\mu} \rangle$	$\langle N \rangle$	$\langle \tilde{\mu} \rangle$	$\langle \hat{\mu} \rangle$
S5 E@H all-sky	2.7	1.2×10^4	0.3	1.2	9.0×10^4	2.7
S5 E@H galactic center	6.2	2.6×10^5	7.2	2.6	5.0×10^4	1.5
S5 all-sky	0.98	260.0	9.8×10^{-6}	0.42	2.0×10^4	9.2×10^{-4}
VSR low-frequency all-sky	1.5	1.2×10^3	3.8×10^{-3}	0.65	5.9×10^3	2.2×10^{-2}
S5 supernova remnant (Cas A)	0.16	3.0	—	6.8×10^{-2}	12.0	—

Table 7.4: Predictions for the expected number of glitches and metric-mismatch at the *highest* spin-down rates for the searches listed in Table 7.1; these are worst cases since we have chosen the highest spin-down rates. We present results both for the number of expected glitches $\langle N \rangle$, the fully-coherent metric-mismatch $\langle \tilde{\mu} \rangle$, and the semi-coherent metric-mismatch $\langle \hat{\mu} \rangle$. Note that the supernova remnants search had no semi-coherent stage and we give the estimate for the search for Cassiopeia A only.

7.6 Conclusions

We have investigated the effects of glitches in the GW signal when searched for using semi-coherent and fully-coherent matched filtering techniques.

In Section 7.3 we confirmed the observation by Espinoza et al. [59] that glitches can be regarded as originating from two distinct populations named Vela-like and normal, with the Vela-like undergoing larger glitches. We then separated the data according to their predicted source population and found fitting formulae for the glitches magnitudes using the spin-down rate $\dot{\nu}$ as a predictor variable. Separating the populations was necessary to avoid overestimating the glitch magnitudes when extrapolating into the parameter space of all-sky gravitational wave searches where few pulsars have been observed. We then used fitting formulae based on data from Espinoza et al. [59] to investigate the rate of glitches.

In Section 7.4 we calculated the metric-mismatch for a glitch with no exponential relation. This was done both for a fully-coherent and semi-coherent searches (Eqn. (7.12) and Eqn. (7.22) respectively) assuming that a single glitch occurs during the search. For each type of search, we provided some simple estimates of the mismatch for the largest glitches seen in the glitch catalogue.

Finally, we transformed our fits for the glitch magnitudes in the EM channel into predictions for the continuous GW channel assuming that $f_s = 2\nu$. This prediction for the glitch magnitude was then used to estimate the mismatch for typical search durations. This predicts that in the initial semi-coherent search, a single glitch will cause a moderate level of mismatch if either the glitch is Vela-like, or normal with the neutron star having a large absolute spin-down rate. If a candidate signal with a glitch does get captured in the semi-coherent stage, we show that, if naively followed up by a single fully-coherent search over the full observation time, it will most likely have a mismatch greater than 10% unless it has a very low absolute spin-down rate. These calculations assume that a single glitch occurred during the search. To complete the picture, we used our fitting formulae for the expected number of glitches to show that for typical search durations there is a reasonable chance that a glitch will occur.

The mismatch estimates potentially overestimate the effect by ignoring the exponential recovery from glitches. If these have short recovery timescales compared with typical observation times, this could reduce the mismatch by a factor of $(1 - Q)^2$, where Q is the glitch recovery parameter. It's unclear exactly what value Q should take for the target population, but we feel this is unlikely, even in the most optimistic scenarios, to nullify the risk.

The levels of mismatch measured here are of concern to both future and past all-sky searches for continuous GWs from neutron stars. If the effect of glitches is ignored, detectable signals could easily be missed due to the presence of a glitch. In a fully-coherent search the presence of a glitch can easily be determined either by including it as a search parameter, or by considering different sections of data. A glitch has a weaker effect on a semi-coherent search and so it is possible the signal will be identified as a candidate, but subsequently lost in the follow-up. We therefore recommend modifying follow-up procedures by introducing a greater number of steps. By studying how the mismatch increases with the coherence time candidates when a glitch can be identified and followed up using a search template which includes a glitch.

Appendix 7.A Bayesian model comparison: test of mixture models

It seems clear by eye that the histogrammed magnitudes of the frequency change in a glitch, $\log_{10} |\delta\nu|$, as shown in Fig. 7.1, exhibit at least two distinct modes, which suggests that it is generated by more than one mechanism. To model this, we will use a Gaussian mixture model (GMM) [64] with N components. This model assumes that the measured data is taken from a population with N sub-populations, each having a Gaussian distribution with separate mean, variance and weight $(\mu_i, \sigma_i^2, \lambda_i)$ where $i \in [1, N]$; note that $\sum_1^N \lambda_i = 1$. Furthermore, we can also allow each of the components to be skewed with a dimensionless skew parameter α_i which can be either positive or negative determining the direction of the skew, or 0, for which there is no skew. Following O'Hagan and Leonard [129] then the probability density function of the i^{th} skewed Gaussian component is

$$f(x; \mu_i, \sigma_i, \alpha_i) = 2\mathcal{N}(x; \mu_i, \sigma_i) \int_{-\infty}^x \mathcal{N}(\alpha_i x; \mu_i, \sigma_i) dx, \quad (7.27)$$

where \mathcal{N} denotes the Gaussian distribution.

Let y_i be the measured values of $\log_{10} |\delta\nu|$ and ϑ be the collection of all model parameters $\{\mu_i, \sigma_i, \alpha_i, \lambda_i\}$. Then the probability density for a GMM with N components is

$$P(y_i | \text{model}, \vartheta) = \sum_{i=1}^N \lambda_i f(y_i; \mu_i, \sigma_i, \alpha_i). \quad (7.28)$$

To compare different choices of N , we will perform a Bayesian model comparison (see §4.2 of Jaynes [86] for an introduction) between each of the mixture models and the simplest hypothesis, a mixture model with $N=1$.

For each model parameter we must specify a prior. We list these in Eqn. (7.29) having defined $\langle y \rangle$, $|y|$, and $\text{std}(y)$ as the average, range, and standard-deviation of the data.

$$\begin{aligned} P(\mu_i) &= \text{Unif}(\langle y \rangle - |y|, \langle y \rangle + |y|), \\ P(\sigma_i) &= \text{Half-Cauchy}(0, \text{std}(y)), \\ P(\lambda_i) &= \text{Unif}(0, 1), \\ P(\alpha_i) &= \mathcal{N}(0, 10 \times \text{std}(y)). \end{aligned} \quad (7.29)$$

For the mean μ_i we use a uniform prior over a range of values containing all data points. For the standard-deviation σ_i , we will use a Half-Cauchy distribution with zero-mean as suggested by Gelman et al. [65]. A large standard-deviation, as compared to the standard deviation of the data itself, provides a weakly informative prior. Instead, we use a standard deviation of $\frac{\text{std}(y)}{2}$ to favour GMM components with small standard-deviations as compared to the data. That is, our prior disfavours models in which any of the components are wide and flat. The prior for λ_i is uniform on $[0, 1]$ and for α_i is normally distributed with zero mean and a wide, weakly-informative standard deviation. The choice of a zero mean favours non-skewed components. Note that, the non-skewed models do not include α_i as a model parameter and the GMM with $N = 1$ does not include λ_i .

We use this choice of prior for the model parameters of each component in the GMM with N components. In this way, models with larger values of N have a larger ‘prior volume’ and hence there is a natural Occam-factor favouring the simpler models with fewer components; this prevents over-fitting.

We will present results for the Bayes factor between a GMM with N components and the simplest model, a GMM with $N = 1$ components. This is computed by

$$\frac{P(\text{model}|\{y_i\})}{P(N=1|\{y_i\})} = \frac{\int_{\theta} P(\{y_i\}|N \text{ GMM})P(\theta)d\theta}{\int_{\vartheta} P(\{y_i\}|N=1 \text{ GMM})P(\vartheta)d\vartheta}. \quad (7.30)$$

We use the *emcee* [62] MCMC algorithm to sample from the posterior and thermodynamic integration to estimate the evidence integrals [67]. In Table 7.5 we provide the \log_{10} of the Bayes factor for several possible models. The Bayes factor between any two of the models given in Table 7.5 can be calculated from their difference.

This table clearly shows that the data is decisive: a Gaussian mixture model with $N \geq 2$ fits the data a great deal better than the simple $N = 1$ GMM. This is unsurprising given the distinct multimodal nature of the data. However, the differences between the other models is more subtle. No single model distinguishes itself by a decisive odds-ratio compared to its neighbouring models. We have checked that these results are robust to small changes in the prior specification.

model	$\log_{10} \left(\frac{P(\text{model} \mathbf{d})}{P(N=1 \text{ GMM} \mathbf{d})} \right)$
2-components	39.12 ± 0.19
2-components (skewed)	41.60 ± 0.21
3-components	42.70 ± 0.23
4-components	44.27 ± 0.24
5-components	44.18 ± 0.22
6-components	43.21 ± 0.22
7-components	42.26 ± 0.22

Table 7.5: Table of the Bayes-Factor for all models considered in this study compared to the simplest $N=1$ GMM. The error is an estimate of the numerical error in the thermodynamic integration.

To help illustrate the differences between some of these models, in Figure 7.14 we plot the probability density for the maximum posterior model parameters found for each model. It is clear from these plots that the $N = 2$ model is not perfect, especially in

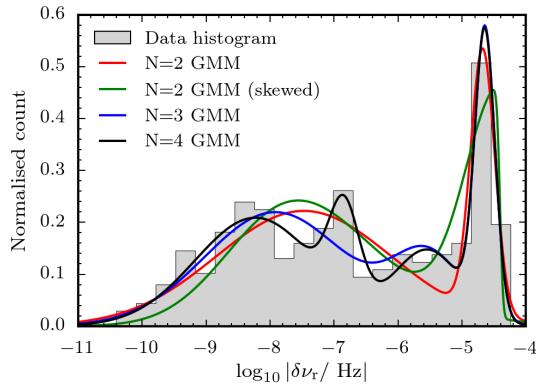


Figure 7.14: The distribution of glitch sizes in frequency along with the predictions for the components of several GMM. The Bayes-factor for these models can be found in Table 7.5.

predicting the number of glitches found in between the two primary subpopulations, around $\log_{10} |\delta\nu/\text{Hz}| = -5.5$; by comparison the $N > 2$ models and the $N = 2$ model which allows for skewness can explain these points and this is reflected in the Bayes factor.

From this analysis it is difficult to decide which model best fits the data. However, what is clear is that simply modelling the data as a GMM with two components with a skew provides a reasonable empirical model. For this reason, in our analysis of the glitch population we will use this model and not any of the models with a greater number of components.

It is important to realise that this comparison is purely empirical, in that the result was not conditioned on a substantive physical model. It would be interesting to include such modelling, this may provide some insight into the appropriateness of the mixture model and the number of components.

Appendix 7.B Linear regression in log-space

In Section 7.3 we perform several linear regressions in log-space in order to calculate power-law fits. This assumes that the observed values $\log(y_i)$ depend on the predictor values $\log(x_i)$ as

$$\log(y_i) = m \log(x_i) + c + \epsilon_i \quad (7.31)$$

where the ϵ_i are independent and identically central normally distributed variables with a standard-deviation σ . In this way, m and c are the linear fit free variables, while σ is a measure of the variability in the observation about this linear fit.

We use a Bayesian linear regression in which we estimate the posterior distributions of all three parameters using a Markov chain Monte Carlo algorithm; for the prior distributions we use non-informative priors and test that these do not induce any bias. In all cases we find the resulting posteriors to be Gaussian and so can take their mean values to get best-fit parameters. The advantage of this method compared to a simple least-squares linear regression is that we also estimate $\langle\sigma\rangle$, the variation about the linear fit. The linear fit can therefore be written as

$$y(x) = \langle m \rangle x + \langle c \rangle \pm \langle \sigma \rangle \quad (7.32)$$

We can then rearrange this equation to give the corresponding power law fit in linear space

$$y(x) = 10^{\langle c \rangle} x^{\langle m \rangle} 10^{\pm \langle \sigma \rangle} \quad (7.33)$$

where the last term gives the variability about the mean. Hence neglecting this term gives the mean.

This is an inherently problematic approach since many functions besides a power law can appear linear in a log-log plot and the assumption of Gaussian error itself may not be valid. Nevertheless, we will still apply this approach since we need only order-of-magnitude estimates and can always check our predictions; we must be clear that the power-law fit gives a good empirical fit, but is not intended to signify any substantive underlying model.

Appendix 7.C Understanding the uncertainty in the predictions

As with all prediction, our estimates carry uncertainties both in the process itself, and in the spread of the observed population. In calculating the expected mismatch, the biggest source of errors come from the variability in the size of glitches in both $\delta\nu$ and $\delta\dot{\nu}$, as shown by the shaded bands plotted in Figure 7.4. Also, calculating the expected number of glitches we have uncertainty from the variability as shown by the shaded bands in Figure 7.6. To understand how these uncertainties may change our belief in

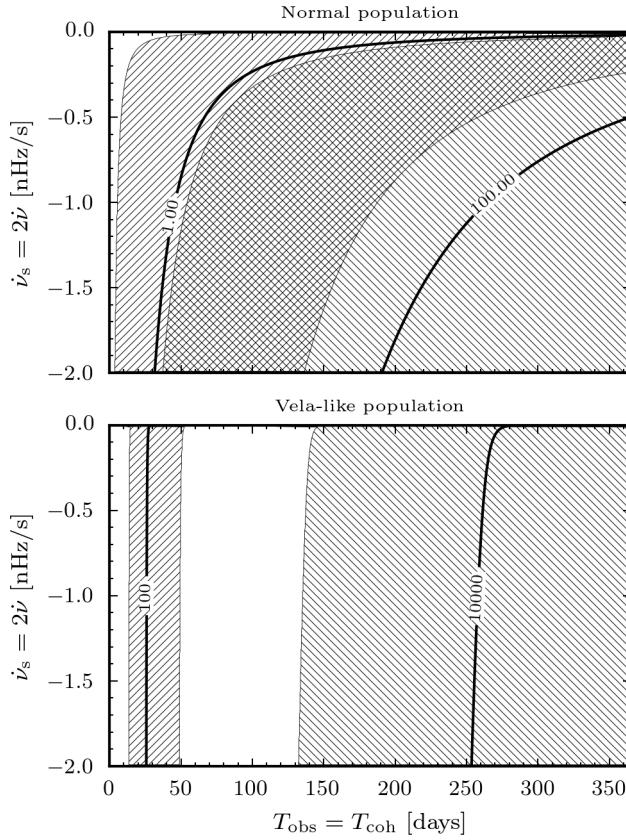


Figure 7.15: A reproduction of the mismatch contours in Figure 7.10 with a reduced number of contours, but showing the variation in predicted mismatch due to the 1σ uncertainty in the predicted distributions as given by the shaded bands in Figure 7.4. The two uncertainty bands overlap, giving rise to the cross-hatched region.

the conclusions, in Figure 7.15 we have repeated the analysis that led to Figure 7.10, with a reduced number of contour lines, and propagated the measure of uncertainty in the glitch sizes. Specifically, we fill contour lines between the uncertainties in the mismatch, given by the error in Eqn. (7.1) and Eqn. (7.2). In Figure 7.16 we repeat the exercise showing the uncertainty on the number of glitches as given by the error in Eqn. (7.3).

These figures illustrate that our uncertainty in exactly where the contour lines sit is large. Nevertheless, even in the best case scenarios large portions of the parameter space remain at risk.

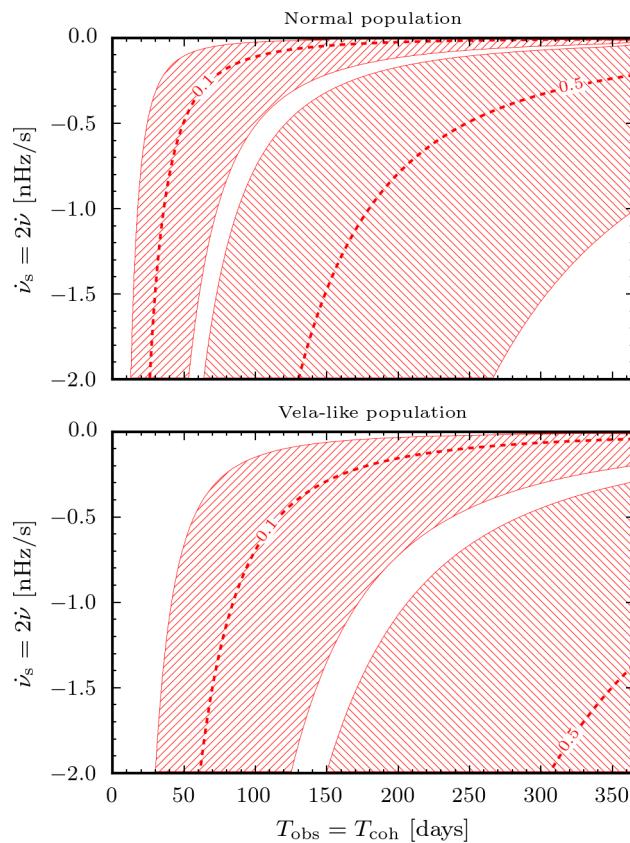


Figure 7.16: A reproduction of the number of glitches in Figure 7.10 with a reduced number of contours, but showing the variation in the expected number of glitches due to the uncertainty in the glitch rate fit of Eqn. (7.3).

Chapter 8

Timing noise in continuous gravitational waves: a numerical study

In this chapter, we will investigate the effect of timing noise on targeted and narrow-band coherent searches for continuous gravitational waves from pulsars. In particular, we will use the Crab ephemeris data (introduced in Section 8.2) to simulate ‘noisy’ gravitational wave signals and study how standard search tools perform when attempting to detect the signals. This chapter contains material reprinted with permission from Ashton et al. [24] as follows: Ashton, G.; Jones, D. I.; Prix, R. Physical Review D, Volume 91, Issue 6 (2015). Copyright (2015) by the American Physical Society.

8.1 Introduction

Timing noise is often represented by structure in the timing residual, which is the difference between the best fit Taylor expansion, typically up to second order in spin-down \ddot{f}_0 , and the observed phase. Timing noise refers specifically to deviations from Taylor expansions that are intrinsic to the pulsar and not to systematic errors such as dispersion in the interstellar medium. Hobbs et al. [81] conducted a wide ranging study on timing noise across the pulsar population. They concluded, amongst other things, that timing noise is ubiquitous and inversely correlated to the age of the pulsar. There already exists measures used to quantify the strength of timing noise such as the Δ_8 value introduced by Arzoumanian et al. [23], the generalisation of the Allan variance [119], the covariance function of the residuals [42], and the red spin noise parameter calculated when fitting for timing noise as part of the pulsar timing model [104]. These do not convert directly into the effect that timing noise may have on continuous GW searches for pulsars. To quantify this, we need to measure the *mismatch* (as defined in Section 6.2) due to timing noise. This is closely related to the loss of signal to noise ratio due to the imperfect matching between the template and signal (which we define explicitly in Section 8.3).

As discussed in Chapter 2.3, a variety of models exist to interpret timing noise, but there is currently no consensus on a single mechanism. However, for the issue of timing noise and continuous GW searches, we only need to consider the relation between the components of the neutron star which produce the EM and GW signals. This was investigated by Jones [88] who identified three possible scenarios. First, the two signals are strongly coupled: the same timing noise will be observed in both. Second, the two signals are loosely coupled: a similar, but different level of timing noise will be observed in both. Third, timing noise exists only in the EM signal, there is no corresponding variations in the GW signal. Of course these are really three cases from a full spectrum of possibilities which could also include the pulsars GW signal being significantly more noisy than its EM signal.

The significance of timing noise will vary between different types of continuous GW searches; these can be divided into targeted, narrow-band, directed, and all-sky searches. *Targeted* searches involve a single known pulsar where an estimate of the spin parameters has been obtained from the EM signal. If we assume that the EM and GW signals are strongly coupled, then we can use a *single-template* targeted search. Under this assumption, when the level of timing noise in the EM signal is small, then a single Taylor expansion is sufficient. If instead the level of timing noise is large, then the EM data can be used to account for it; this is done by applying an adapted matched filtering phase-model that closely follows the observed EM phase model [137]. If instead we assume that the EM and GW signals are loosely coupled, then we should perform a *narrow-band* search in a small area of parameter space. These narrow-band searches aim to allow for small frequency offsets between the EM and GW signals, such as could be caused by free precession, or a finite coupling time between the two components of the neutron star [11]. *Directed* searches look for non-pulsing neutron stars predicted by other means such as at the centre of the supernova remnant Cassiopeia A. An *all-sky* search involves searching over the entire sky for unknown pulsars. For both directed and all-sky searches the lack of EM data necessitates wide bands in the frequency and its derivatives. For fully coherent matched filtering methods these searches can rapidly become computationally prohibitive. To circumvent this, semi-coherent search techniques are used that incoherently combine short fully-coherent sections of data [8]; these will be less sensitive to timing noise. Nevertheless, semi-coherent searches ultimately need to be followed up by targeted fully coherent searches, for which timing noise may be an issue.

For the properties of the GW signal, the most general case is that it will exhibit some timing noise, but it could be different to the timing noise observed in the EM signal. Until a detection is made, we can only make assumptions about how the two are correlated. To probe these assumptions, we will define two special cases corresponding to different sorts of errors in a GW search:

- **Special Case 1:** Timing noise, exactly like that in the EM signal, exists in the GW signal but is not included in the template. This will result in a loss of signal to noise ratio for searches which assumed that timing noise was negligible. The error potentially affects the narrow-band, directed, and all-sky searches since the

level of timing noise is unknown. The single template targeted searches will not be effected since they either check that the level of timing noise is negligible, or correct for it using an adaptive phase model.

- **Special Case 2:** Timing noise is included in the template but does not exist in the signal. This will result in a loss of signal to noise ratio for single-template targeted searches that account for timing noise using an adapted phase model (for example Abbott et al. [11]). Instead, these searches will now erroneously introduce timing noise into the template while the signal will be a smooth Taylor expansion.

In this chapter, we will mimic narrow-band and single-template searches to directly simulate special case 1. Specifically, we will inject a fake GW signal which contains a realisation of timing noise, and recover it using templates based on a single global Taylor series. This tests the scenarios in which the timing noise in the GW signal is either exactly coupled to the EM signal, or they are at least similar. However, this also quantifies special case 2 since the signal and template are interchangeable in matched filtering methods. That is, timing noise in the signal, but not in the template is equivalent to timing noise in the template, but not the signal.

While all known pulsars are potential GW sources, young pulsars are the most promising due to their large spin-downs (see Abbott et al. [13] for a review). However, it was found by Hobbs et al. [81] that the amount of timing noise is correlated with the spin-down magnitude. This motivated us to study the effect of timing noise on GW searches for neutron stars with large spin-downs.

The realisation of timing noise we will use to investigate timing noise in GW pulsar sources is based on the young Crab pulsar. The Crab is a potentially detectable source of gravitational waves due to its high spin-down rate and it has the highest spin-down upper limit compared to the LIGO noise floor [11]. The EM signal from the Crab is well documented (see Section 8.2) and contains exceptional levels of timing noise: it was estimated by Jones [88] that such levels of timing noise in the GW signal may cause an issue for current searches.

Several targeted searches have already been performed for GWs from the Crab pulsar. A single-template search for GWs from the Crab pulsar was performed on data collected during the LIGO S5 science run [11]. This search used the Crab ephemeris and an adapted phase model to account for timing noise. In addition to this single-template search, a narrow-band search for signals from the Crab was also performed by Abbott et al. [11] on the S5 data. Another narrow-band search for the Crab was carried out using data from the VIRGO VSR4 science run along with a search for the Vela pulsar [5].

The structure of this chapter is as follows. In Section 8.2, we describe the observational data available from the Crab ephemeris and discuss its relation to GW searches. In Section 8.3, we describe the signal injection and recovery method. Results from this method are presented in Section 8.4: we begin by considering the effect timing noise has on narrow-band searches, then we consider the mismatch on stretches of data for which narrow-band searches have been performed; we further investigate how the mismatch

depends upon epoch; and we finally examine how the mismatch depends on the duration of observation. Finally, we summarise our results in Section 8.5.

8.2 Timing noise as described by the Crab ephemeris

The monthly Crab ephemeris [113] provides the phase evolution of the EM signal between 1982 and the present and can be found at <http://www.jb.man.ac.uk/pulsar/crab.html>. It is unlike most timing data for pulsars where a timing model consists of the model parameters (position, spin-down, etc.) given at a single reference time. For the Crab ephemeris, each monthly update consists of the frequency and spin-down coefficients along with a reference time coinciding with the time of arrival (TOA) of a pulse at the solar system barycentre. The coefficients are calculated by least-squares fitting of a Taylor expansion to the TOAs. The reference time for each month is chosen as the TOA of the pulse closest to the mid-point; this is done to minimise the average phase error of the local Taylor expansion. The period of a month is short enough such that these coefficients and the Taylor expansion, Eqn. (1.2) track the rotational phase during the month. The Crab ephemeris tracks the rotational frequency and frequency derivative of the Crab pulsar (ν and $\dot{\nu}$). In this chapter, we will consider searches for gravitational wave emission from non-axisymmetric neutron stars at twice this frequency $f = 2\nu$; hereafter all frequencies f and spin-downs \dot{f} refer to the pulsar's GW emission.

The Crab ephemeris gives a distinct picture of the variations due to timing noise superimposed on the monotonic spin-down. To illustrate how this manifests itself, Figure 8.1 depicts the frequency evolution in two adjacent months. Notice that a discontinuity occurs at the interface between months. Such discontinuities will occur in the spin-down, frequency, and phase; timing noise can then be described by the magnitude of these jumps. From the Crab ephemeris it can be shown that the distribution of jumps in phase, frequency and spin-down appear to follow standard normal distributions. This is consistent with timing noise models consisting of a large number of small unresolved events accumulating over a month (e.g. the models considered by Cordes and Greenstein [47]). The Crab ephemeris therefore provides an alternative way to discuss timing noise (rather than in terms of structure in the timing residual), we investigate this further in Section 9.5.1.

Timing noise is usually depicted by structure in the phase residuals calculated by removing the best fit Taylor expansion to the phase from the real phase evolution. A best fit Taylor expansion consists of a single set of coefficients f_0 , \dot{f}_0 , and \ddot{f}_0 valid over the *entire* observation period. To make this distinct from the *local* Taylor expansions describing the evolution in each month, this will be referred to as the *global template*. In Figure 8.1 we see that if the discontinuity is non-zero, then it is impossible for any global Taylor expansion template to exactly match the local signal in both months. The phase residual, and hence timing noise, results from the inability to match a single global template to the non-smooth signal.

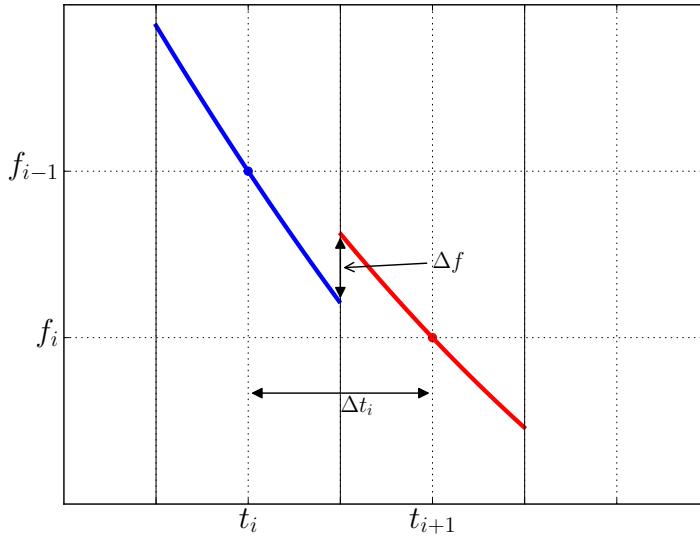


Figure 8.1: Illustration of the jumps between ‘local’ per-month signals in frequency space, defining the frequency jump Δf . This depiction amplifies the order of magnitude of Δf in order to highlight the timing noise: for the jumps in the ephemeris, Δf is several orders of magnitude smaller than the change in frequency due to spin-down alone.

In this chapter, we aim to quantify the significance of timing noise in GW searches by generating signals from the Crab ephemeris. This is an empirical description of timing noise and so we make no assumptions on the underlying astrophysical model.

8.3 Method

We now describe the method to quantify the effect of timing noise on continuous GW searches for signals from isolated pulsars. To be relevant to current pulsar continuous GW search methods, we will base our method on the narrow-band searches of Abbott et al. [11] and Aasi et al. [5]. The results can be interpreted as measuring the consequence of special case 1 on narrow-band and single-template searches; that is we assume the GW signal has a similar level of timing noise as the EM signal and search using global Taylor expansion templates. For this study, a single-template search refers to a single Taylor expansion template and not the adapted phase model proposed by Pitkin and Woan [137].

We begin by generating a GW signal emulating timing noise using the Crab ephemeris. This is done by stringing together month-long smooth Taylor expansion signals. Each month uses the corresponding month from the Crab ephemeris for the Taylor expansion coefficients. The ‘jumps’ at the interface between months constitutes the timing noise. In more detail:

1. From the ephemeris select a period of data consisting of the reference times t_i , frequency f_i , and spin-down \dot{f}_i for each month i .
2. Generate the phase as a function of time from the data and then fit a global Taylor expansion up to \ddot{f} for the whole observation time. The fit results in a set of interpolated coefficients $[f_0, \dot{f}_0, \ddot{f}_0]$ at a global reference time halfway through the data. These coefficients are used to centre the narrow-band search parameters.
3. We supplement the local monthly data $\{t_i, f_i, \dot{f}_i\}$ with the fixed value of \ddot{f}_0 calculated in the previous step. The phase of the GW signal is always zero at each monthly reference time t_i of the ephemeris, which by construction coincides with a pulse arrival time.

In this process we have assumed that a fixed value of \ddot{f}_0 is sufficient. This can be justified by considering the next term in the Taylor expansion and typical values of $\ddot{f} \sim 10^{-30} \text{ Hz/s}^3$. Over typical search durations ~ 1 year, this term contributes less than a radian to the phase (which is less than typical contributions from timing noise in the frequency and spin-down rate), and it can therefore be safely neglected.

We use the **LALSuite** [106] gravitational-wave analysis routines to generate data with a fake GW signal; for these experiments we work without any simulated detector noise. The standard tool to generate fake GW signals uses single Taylor expansion models. Therefore, to include timing noise in the signal we do the following.

4. Into the data which we search for a signal, we inject each month-long Taylor expansion generated from the Crab ephemeris lasting for only the duration of that month. This method creates a fake GW signal, lasting several months, which includes timing noise corresponding to the monthly ephemeris.

Once we have produced data, we then use **LALSuite** tools to recover the signal using the \mathcal{F} -statistic [85]. This is a matched filtering method in which the output of the detector is compared to a signal template (see Prix [140] for more details).

Two types of searches are performed: a single template search at the interpolated coefficients $[f_0, \dot{f}_0, \ddot{f}_0]$ and a narrow-band search in f and \dot{f} centred on the interpolated coefficients. These searches were found to be sufficient to find the signal to within a reasonable mismatch, so more sophisticated methods were not required.

The narrow-band consists of a grid of points in f and \dot{f} . As found by Abbott et al. [11] we find searching over \ddot{f}_0 to be unnecessary for this experiment and so it is kept fixed using the value found in step 2 above. The grid spacing is parameterised by m , the one-dimensional maximal mismatch between two adjacent Taylor expansion templates. From Aasi et al. [3] the corresponding grid spacing is given by

$$df = \frac{\sqrt{12m}}{\pi T_{\text{obs}}} \quad d\dot{f} = \frac{\sqrt{720m}}{\pi T_{\text{obs}}^2}, \quad (8.1)$$

where T_{obs} labels the observation time.

For the single-template and at each grid point in the narrow-band search, we measure the squared SNR value ρ^2 . In order to quantify the relative loss compared to the perfectly phase-matched squared SNR ρ_s^2 , we define the mismatch in the usual way (e.g. see Prix [139]) as

$$\mu = \frac{\rho_s^2 - \rho^2}{\rho_s^2}. \quad (8.2)$$

It is well known (e.g. see Prix [140]) that the SNR for a perfectly phase-matched signal is independent of the signal phase evolution. Therefore, in the absence of timing noise the measured value of ρ^2 can reach the maximum value of ρ_s^2 , and the mismatch therefore vanishes in that template. In the presence of timing noise, even the best-matching template will suffer some mismatch, and this effect will increase with the level of timing noise.

In the single-template search, we measure a single mismatch value. The single-template search can also be interpreted to quantify the error made in special case 2, when the template is adapted to account for EM timing noise but none exists in the GW signal. We can think of the narrow-band search as repeating the single-template search over a grid of points; this allows us two degrees of freedom, corresponding to the frequency and spin-down parameters, over which to minimise the mismatch. The grid point with the minimum mismatch, which we denote by μ_{\min} , is the best candidate and will be used to quantify the success of the search. Because the narrow-band can minimise the mismatch, μ_{\min} must always be equal or smaller than the mismatch in the single-template search.

8.4 Results

8.4.1 The effect of timing noise on narrow-band searches

We begin by describing how timing noise degrades a narrow-band search. This is done by comparing the result for a signal containing no timing noise with a signal generated from the Crab ephemeris between MJD 45150 and 45668. This period holds no special significance and is used simply to demonstrate the essential features of a signal containing timing noise.

In Figure 8.2 we show the mismatch as a function of parameter space offset for (a) a signal without timing noise, and (b) a signal containing timing noise. The signal without timing noise is injected at the interpolated coefficients $[f_0, \dot{f}_0, \ddot{f}_0]$. Therefore, we find the minimum mismatch with $\mu_{\min} = 0$ at exactly the centre of the grid and the iso-mismatch contours in the local neighbourhood around the origin are well described by ellipses (e.g. see Prix [139]). For the signal with timing noise (b) we notice two distinctive effects: the minimum achievable mismatch μ_{\min} is non-zero, and the iso-mismatch contours around the best-match template are more irregular and less well described by ellipses. In the following we will quantify the effect of timing noise by considering only the location and value of the minimum mismatch grid point in the narrow-band search.

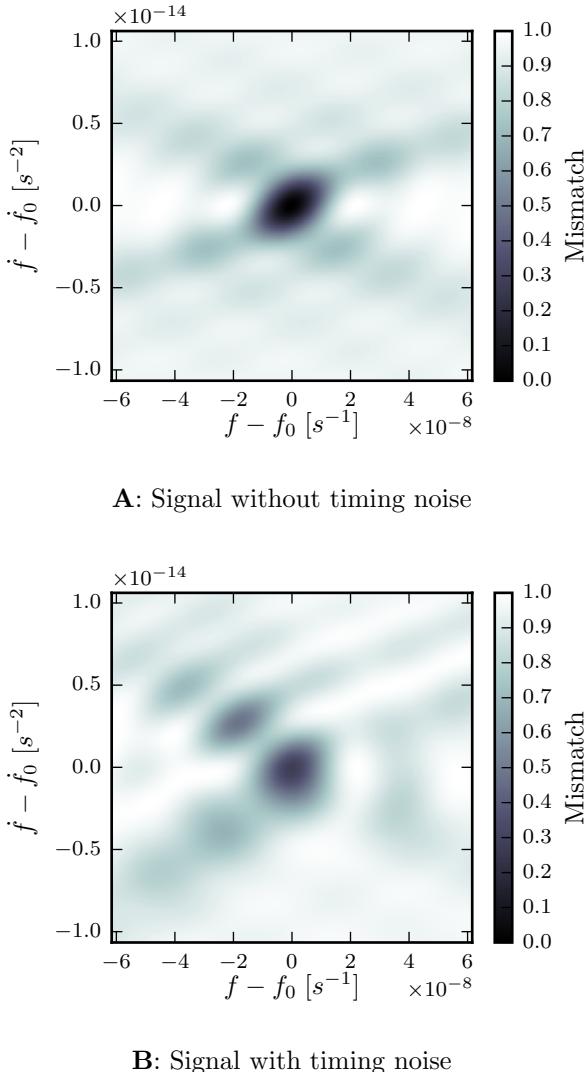


Figure 8.2: In Figure **A** we show the mismatch as a function of parameter space for a signal without timing noise. The injected signal has parameters $[f_0, \dot{f}_0, \ddot{f}_0]$. As a result, the mismatch has a minimum at this point. This can be compared with Figure **B** showing the mismatch from a signal including timing noise. The signal is generated from the Crab ephemeris between MJD 45150 and 45668.

8.4.2 Results relevant to recent narrow-band searches

First we consider two particular periods of the Crab ephemeris corresponding to recent narrow-band searches for the Crab: the LIGO S5 period [11] and the VIRGO VSR4 period [5]. The mismatch in the single-template and the minimum mismatch for the narrow-band searches during both periods are listed in Table 8.1. For these periods timing noise is found to produce a mismatch of $\approx 1\%$. As expected, the narrow-band mismatch is smaller than the single-template search. The fractional difference between the two searches is relatively small.

Provided that the timing noise observed in the GW signal is at the same level (or less) as that observed in the EM signal, this result signifies that the recent LIGO and VIRGO narrow-band searches would not suffer significantly from the effects of timing noise.

In addition to producing a mismatch, timing noise may result in the best candidate being found at some distance from the centre of the narrow-band search. However, we find that the distance from the centre of the grid is small when compared to the grid spacing used in actual narrow-band searches such as that used for the S5 and VSR4 searches. For the S5 period narrow-band search, we find that the minimum mismatch was a fraction ~ 0.01 of the grid spacing used in the Abbott et al. [11] search. At the resolutions used in real narrow-band searches, the effects of timing noise on the location of the minimum mismatch will not be evident.

	Dates	Single template	Narrow band
	MJD	μ	μ_{\min}
S5	53673 - 53977	0.00968	0.00933
VSR4	55681 - 55839	0.00659	0.00584

Table 8.1: Measurements of the mismatch during the S5 and VSR4 narrow-band search periods.

Figure 8.3 shows the convergence of the measured best mismatch μ_{\min} for the narrow-band search over the S5 period with the value of m . This demonstrates that the non-zero values of μ_{\min} given in Table 8.1 are not the result of grid coarseness. For signals without timing noise, the measured best mismatch μ_{\min} will have a minimum of $\sim m$ when the putative signal is located halfway between grid points. In the limit of $m \rightarrow 0$ we then expect the measured mismatch to tend to zero. Instead, for a signal with timing noise we observe a plateau after some initial reduction. This indicates that the grid is now *fully* resolving the variations due to timing noise.

8.4.3 Minimum mismatch as a function of the observation epoch

We will now investigate how the best mismatch μ_{\min} varies as a function of the observation epoch. We only show the narrow-band search, as the results were found to be very similar for the single-template search. The method consists of measuring the mismatch μ_{\min} in a 6-month window, which is shifted in 1 month intervals over all the available ephemeris data. The observation time of 6 months is chosen to be similar to typical GW search durations. We are restricted to multiples of 1 month by the frequency of updates to the Crab ephemeris.

Timing noise is not the only variability in the spin-down of pulsars - they can also undergo sudden increases in rotation frequency known as *glitches* (for an introduction, see Section 2.1). The Crab frequently glitches and these are catalogued by Espinoza et al. [59] and available at <http://www.jb.man.ac.uk/pulsar/glitches.html>. The mechanism which causes a glitch is not well understood and may involve unpredictable variations in the GW signal. As a result, targeted GW searches either avoid periods with known glitches [11], or allow for an arbitrary jump in gravitational wave phase

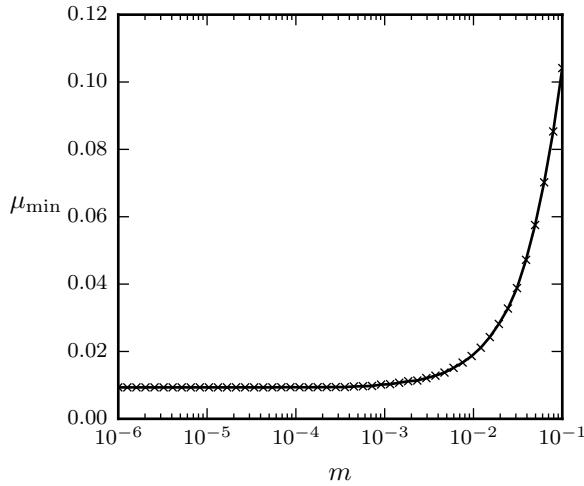


Figure 8.3: Measured best mismatch μ_{\min} as a function of grid spacing parameter m (see Eqn. (8.1)), for the Crab pulsar over the S5 period. This demonstrates that μ_{\min} plateaus at a nonzero mismatch suggesting we are resolving a mismatch due to timing noise instead of the effect of finite grid resolution.

at the time of the glitch [13]. For this chapter, we will ignore the complicating factor introduced by glitches (which was considered in Chapter 7) and consider separately the effect of timing noise. We do this by omitting windows which include glitches from the search by using the aforementioned glitch catalogue.

We begin by searching in a small 40×40 grid in frequency and spin-down, with a fixed grid space mismatch of $m = 1 \times 10^{-5}$, and the grid spacing as defined in Eqn. (8.1). It is possible that the minimum mismatch is found at the edge of the narrow-band grid; such candidates are not true local minima in the mismatch. If this is the case, the search is repeated with an increasingly larger grid size, but the same fixed grid spacing. This process continues until we find a minimum mismatch which is not at the edge of the grid.

Figure 8.4 shows the measured minimum mismatch in the narrow-band search for a sliding 6-month window at the centre of the observation time. The mismatch due to timing noise is the low level noise occurring in between glitches. Greater mismatches are observed in the post-glitch periods; this is expected as the relaxation time after glitches for the Crab is of the order 1 month [113]. We note the presence of an anomalous period of large mismatch for all windows that include the ephemeris time MJD 55362. The cause for this is unclear from the available data, but it may be caused either by a measurement error or a small undetected glitch. In general, we find that the level of mismatch due to timing noise is between $\mu_{\min} \sim 10^{-3} - 10^{-2}$ for these 6-month searches.

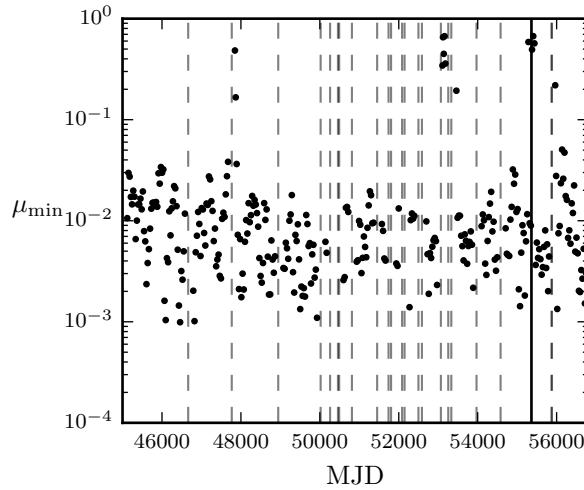


Figure 8.4: Minimum mismatch μ_{\min} found in 6-month sliding window searches as a function of epoch at the centre of window. Vertical dashed lines indicate glitch events as described by Espinoza et al. [59]. The solid vertical line indicates the date MJD 55362, a period of anomalously large mismatch.

8.4.4 Averaged minimum mismatch as a function of the observation duration

We can study the averaged behaviour of the mismatch μ_{\min} as a function of time by varying the size of the sliding window in the previous section. This was done for both the narrow-band and single-template searches; the mismatch from the narrow-band search was found to be a fraction $\lesssim 0.1$ smaller on average than the single-template search. We therefore will only present results from the narrow-band search. The shortest possible window ~ 6 months is restricted by the number of points needed to generate a fit to the phase. Setting the upper limit at ~ 17 months retains a statistically meaningful number of points to average over. Having obtained the data from all sliding window sizes in this range we want to analyse the average behaviour as a function of the observation time. Before doing this we filter results in the following ways:

- We do not consider any windows that include or are bounded by glitch events
- Windows including the anomalous epoch MJD 55362 are omitted. We wish to study the fluctuations due to timing noise, and this period is either an unidentified glitch, or another highly unusual and unrepresentative form of timing noise
- While each entry of the ephemeris is on average valid over a whole month, some months were truncated due to glitches. The sliding window, which works on a fixed number of entries of the ephemeris will occasionally be shorter than average. To ensure we are averaging over windows of a similar length we omit windows for which the observation time differs from the average by 2 weeks.

In Figure 8.5 we plot the averaged minimum mismatch $\langle \mu_{\min} \rangle$ as a function of observation time. This indicates a growth of $\langle \mu_{\min} \rangle$ with observation time resembling a power law.

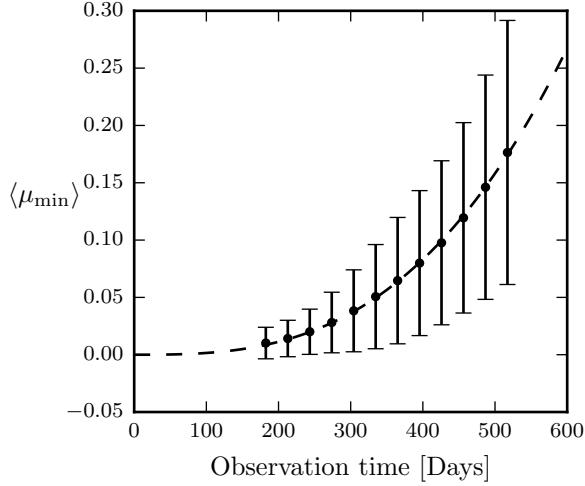


Figure 8.5: Averaging the mismatch for sliding window searches and varying the observation times. The points give the mean while the bars correspond to one standard deviation.

To quantify the growth of the mismatch, we perform a least-squares fitting to a power law. Fitting the expression:

$$\langle \mu_{\min} \rangle_{\text{fit}} = \kappa \left(\frac{T_{\text{obs}}}{1 \text{ sec}} \right)^n, \quad (8.3)$$

we find the best fit parameters

$$\kappa = 1.5 \pm 0.8 \times 10^{-23} \quad (8.4)$$

$$n = 2.88 \pm 0.030. \quad (8.5)$$

In Chapter 9 we will return to this and try to analyse if this fit can help to constrain the source of timing noise.

For perfectly matched signals the squared SNR increases linearly [140] with observation time. This suggests that longer observation times yield a greater likelihood of detection. The power law fit with $n > 1$ implies that the average mismatch from Crab timing noise grows faster than the squared SNR. Gains in SNR from longer observation time will therefore eventually be outweighed by the increasing mismatch from timing noise. To estimate when this may occur, we can rearrange Eqn. (8.2) to give

$$\rho^2 = \rho_s^2 (1 - \mu_{\min}). \quad (8.6)$$

Substituting the time dependencies for the perfectly matched SNR and the averaged mismatch we have

$$\rho^2 \propto T_{\text{obs}} - \kappa T_{\text{obs}}^{n+1}. \quad (8.7)$$

Differentiating and solving for T_{obs} yields an expression for the observation time (in seconds) beyond which the ρ^2 value of a signal containing timing noise starts to decrease

$$T_{\text{obs}} = \left(\frac{1}{\kappa(1+n)} \right)^{1/n}. \quad (8.8)$$

For the fit values from Eqn. (8.5), this yields a critical observation time of $T_{\text{obs}} \approx 600$ days after which the mismatch exceeds $\langle \mu_{\text{min}} \rangle \approx 0.25$. In this case it is no longer true that further increases in observation time will yield greater detectability.

Jones [88] estimated the maximum time the signal and template would remain coherent given a random walk in frequency. A crude method used a phase residual of 1 rad for the decoherence criteria. For the Crab, this estimates the decoherence time at 200 days. We can improve upon this result by setting a mismatch of 0.1 as the decoherence criteria; using the fit to the averaged mismatch this gives us a decoherence time of $T_{\text{obs}} \approx 400$ days.

The growth of mismatch as a power law is suggestive of random walk timing noise models for which, as shown in Section 2.3.1, the root-mean-square phase residual also grows as a power law. The power to which the phase residual grows depends on the type of random walk: [33, 47] classified the Crab as undergoing frequency-like noise. However, to compare with Eqn. (8.5), such a scaling in the phase residual must first be converted into a mismatch, which will depend on the search method, before a comparison can be made. We will return to this in Chapter 9 where we directly calculate the mismatch due to a random walk and update Figure 8.5 with this prediction.

8.5 Conclusions

We have used observational data on the Crab pulsar to characterise the possible effects of timing noise on coherent targeted single-template and narrow-band continuous gravitational-wave searches for pulsars. This was done by generating fake signals based on the Crab ephemeris data and searching for them using templates without timing noise. Our analysis clarifies the impact for current searches; accordingly, our methods mimic those used by Abbott et al. [11] and Aasi et al. [5].

Our principal result is summarised by Figure 8.5: when considering the average mismatch as a function of observation time, we find that the averaged mismatch grows as a power law. In addition to this, we found two interesting aspects when considering the data without averaging over the epoch:

Firstly, for the S5 and VSR4 narrow-band searches, if the timing noise in the GW signal from the Crab is at a similar level (or lower) to that in the EM signal, then we find it will only have a small ($\approx 1\%$) effect on the measured squared SNR of the putative signal. We found the mismatch in single-template searches to be only fractionally larger than the narrow-band searches. This also suggests phase-adapted searches would not be significantly affected if the signal does not contain timing noise.

Secondly, searching over all available Crab data with a 6-month window, we looked at the mismatch as a function of observation epoch. Post glitch periods tend to admit significant levels of mismatch; this is expected due to the exponential recovery from the glitch. (We also discovered a period around MJD 55362 which has a large mismatch and is not connected to a known glitch). The narrow-band and single-template searches performed similarly in this and subsequent tests. Typically, the mismatch due to timing noise for 6-month searches was found to be between 10^{-3} and 10^{-2} .

The scope of this work can be extended to directed and all-sky searches, which target young rapidly spinning down stars which may emit the strongest GWs. These stars are also known to exhibit the highest levels of timing noise and glitch frequently. Crucially the lack of EM data means we cannot be certain a glitch does not occur during the observation and we cannot account for timing noise in the signal. In future work we would like to quantify both these effects and estimate safe upper limits for the search durations.

Chapter 9

Timing noise in continuous gravitational waves: random walk models

It's unclear at this time what causes timing noise, but from Chapter 8 we know that it may pose a problem for continuous GW searches if the phase evolution is affected by timing noise and this is not included in the search templates. In addition to searches for isolated neutron stars, many searches [see for example 4, 102, 124] have been performed for GWs from low-mass X-ray binary systems (LMXBs). These searches experience similar difficulties due to a stronger form of timing noise known as ‘spin-wandering’. This issue was investigated by Watts et al. [165] who quantified the effect for a spin-up of $\dot{\nu}_s$ by defining a decoherence time T_{decoh} , such that $T_{\text{decoh}}^2 \dot{\nu}_s = 1$, after which the Taylor expansion can no longer track the phase. They then estimated a worst-case decoherence time by using the maximal spin-up rate due to the accretion torque and found that for some sources such as the LMXB Scorpius X-1, the decoherence time can be short as ~ 1 week.

In this chapter, we will present some preliminary calculations and results related to modelling timing noise in a continuous GW using a random walk model. Unlike the results of Chapter 8, which used an empirical description of timing noise given by the Crab ephemeris, the results derived here can be applied to any search in which it is thought the signal may undergo a random walk. In this sense, it is equivalent to Chapter 7 in that the ultimate aim is to estimate the risk faced by various searches by inferring from the observed pulsar population. As given here, the task is incomplete and so we will present our calculations and some preliminary results. We will study the effect in the context of searches which assume a smooth Taylor expansion template; advanced methods such as the ‘loosely-coherent’ search introduced by Dergachev [53] could provide a way to search for noisy signals without a loss of signal to noise ratio.

Recent observations by Hobbs et al. [81] suggest that a random walk model does not capture the physics of timing noise in isolated radio pulsars and hence is not a useful way to infer neutron star physics. Nevertheless, the random walk model remains a practical

empirical model; in this section then, we use it as such without requiring it to have any deeper substantive meaning for what causes timing noise. In the same way, the random walk model can also be applied to LMXBs where the amount of spin-wandering could be inferred from fluctuations in the luminosity. From this, we would like in the future to update the estimates by Watts et al. [165] to estimate the mismatch for searches for continuous gravitational waves from LMXBs.

This chapter is organised in the following way. In Section 9.1 we will define the random walk as a piecewise Taylor expansion in which the difference between the signal and template is initially zero. We will then calculate the mismatch for this simple description in Section 9.2. However, setting the difference between the signal and template to zero initially does not minimise the mismatch, so in Section 9.3 we will discuss a pragmatic method to estimate the minimised mismatch; for each predicted mismatch we verify our calculations using Monte-Carlo like numerical simulations. In Section 9.4 we relate this description of a random walk to the compound Poisson process random walk usually discussed in the literature. Finally, in Section 9.5 we begin by discussing the data from the Crab ephemeris (first introduced in Section 8.2) in the context of a random walk. We then go on to apply the results found earlier in the chapter to predict how the mismatch in the Crab depends on the observation time assuming that it undergoes a random walk in the frequency; this is then compared to the results found in Chapter 8.

9.1 Defining a random walk

To calculate the fully-coherent mismatch, we will model the random walk as a zero-mean Gaussian walk in the phase, frequency, and spin-down which jumps regularly at N_{sd} fixed time intervals ΔT , such that the total observation time is $T_{\text{obs}} = N_{\text{sd}}\Delta T$. This allows us to write the signal as a piecewise Taylor expansion with N_{sd} subdomains. Choosing fixed time intervals appears to introduce an additional timescale not usually present in random walk models. However, as shown later in Section 9.4 this is consistent with a large number of unresolved events which are measured over the fixed timescale.

9.1.1 Initial definitions

In this model, we choose the spin-down rate to be the highest order term which undergoes a random walk. Recalling that the parameter space offset $\Delta \dot{f}_i$ (defined in Eqn. (6.9)) is the difference between the signal and template in the i^{th} subdomain, we may define the jump in this difference between the i and $i - 1$ subdomains as $d\dot{f}_i$, such that

$$\Delta \dot{f}_i - \Delta \dot{f}_{i-1} = d\dot{f}_i \sim \mathcal{N}(0, \sigma_{\dot{f}}^2), \quad (9.1)$$

where \mathcal{N} denotes the normal distribution and we have defined $\sigma_{\dot{f}}^2$ as the standard-deviation of the step sizes in the spin-down rate. Rearranging this gives an expression

for the offset in the i^{th} subdomain. By induction we can also write down the $i - 1$ term

$$\Delta\dot{f}_i = d\dot{f}_i + \Delta\dot{f}_{i-1}, \quad (9.2)$$

$$\Delta\dot{f}_{i-1} = d\dot{f}_{i-1} + \Delta\dot{f}_{i-2}. \quad (9.3)$$

We set the initial difference between the signal and template as zero such that $\Delta\dot{f}_0 = 0$, that is we start each random walk from the origin (we will return to this point in Section 9.3). Then as each step proceeds from the previous step, we have that

$$\Delta\dot{f}_i = \sum_{j=1}^i d\dot{f}_j. \quad (9.4)$$

To illustrate this, in Figure 9.1 we plot an example of a random walk in the spin-down as given by Eqn. (9.4).

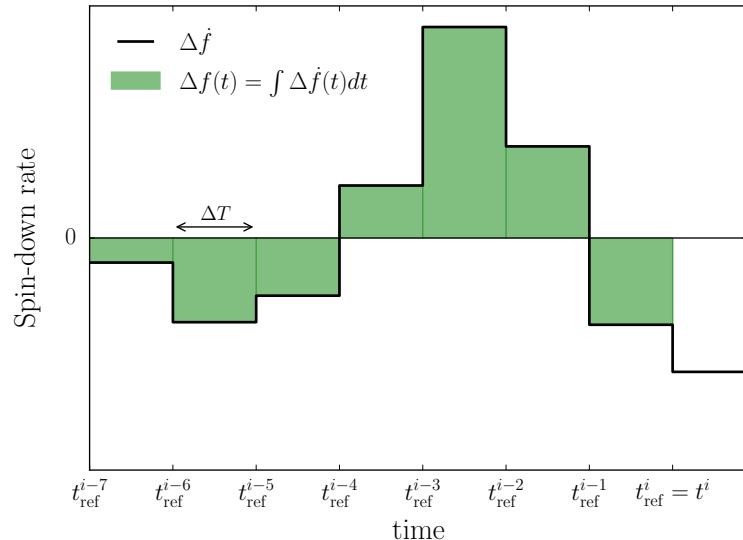


Figure 9.1: An example of a random walk in the spin-down rate, Eqn. (9.4). The filled green blocks indicate the summation defined in Eqn. (9.5) required to calculate the induced change in frequency at t^i due to the random walk in spin-down rate.

If we want to model a random walk in the phase, frequency, and spin-down rate concurrently, then we must consider the effect that a random walk in spin-down will have on the frequency and phase. For example, if we increase the spin-down rate for a period of time, then we would expect the frequency to decrease at a greater rate during this period. In our discrete model, it is not possible to dynamically change the frequency during a single subdomain. However, we can approximate this by updating the frequency in the next subdomain with the induced frequency offset due to the spin-down in all the previous subdomains. This must be done for the induced effect from the random walk in spin-down rate on the phase and frequency, and for the induced effect from the random

walk in frequency on the phase. There is no induced effect for the spin-down rate: the effect only propagates to lower order terms.

Because the random walk is discrete and modelled as constant in any given subdomain, we can calculate the offset in the lower order terms from a Taylor expansion. The total offset at the i^{th} reference time is then given by the summation of the offset caused by all higher order terms up to that reference time. We can choose the reference times arbitrarily, but setting each to start at the beginning of the subdomain simplifies the calculation. For the frequency offset induced by the spin-down, we have

$$\Delta f_i = \sum_{j=1}^{i-1} \Delta \dot{f}_j \Delta T. \quad (9.5)$$

This can be thought of as the integration of the spin-down up to the i^{th} reference time and is illustrated by the green blocks in Figure 9.1.

We want to consider random walks in all three parameters so we now add in a random walk in frequency. Each step is independent of the induced effect from the spin-down and is given by $df_i \sim \mathcal{N}(0, \sigma_f^2)$. The two effects will sum linearly such that the frequency offset is

$$\Delta f_i = \sum_{j=1}^i df_j + \sum_{j=1}^{i-1} \Delta \dot{f}_j \Delta T. \quad (9.6)$$

By a similar process we can calculate the induced effect of the frequency and spin-down on the phase. Including the random walk in the phase for which $d\phi_i \sim \mathcal{N}(0, \sigma_\phi^2)$, the phase offset is given by

$$\Delta \phi_i = \sum_{j=1}^i d\phi_j + 2\pi \left(\sum_{j=1}^{i-1} \Delta f_j \Delta T + \frac{1}{2} \sum_{j=1}^{i-1} \Delta \dot{f}_j \Delta T^2 \right). \quad (9.7)$$

9.1.2 Writing the parameter offsets in terms of normal distributions

Eqn (9.6) and Eqn. (9.7) give the difference between the signal and template as functions of the random walks in higher order parameters. In order to calculate statistical values, we now write these in terms of the normal distributions from which the random walks are constructed. Substituting Eqn. (9.4) into Eqn. (9.6) and using the identity defined in Eqn. (9.76) of Appendix 9.A, we have

$$\Delta f_i = \sum_{j=1}^i df_j + \sum_{j=1}^{i-1} \sum_{k=1}^j d\dot{f}_k \Delta T, \quad (9.8)$$

$$= \sum_{j=1}^i df_j + \sum_{j=1}^{i-1} (i-j) d\dot{f}_j \Delta T. \quad (9.9)$$

Similarly, substituting this equation into Eqn. (9.7) and using both Eqn. (9.76) and Eqn. (9.80) from Appendix 9.A, we have

$$\begin{aligned}
 \Delta\phi_i &= \sum_{j=1}^i d\phi_j + 2\pi \left(\sum_{j=1}^{i-1} \Delta f_j \Delta T + \frac{1}{2} \sum_{j=1}^{i-1} \Delta \dot{f}_j \Delta T^2 \right) \\
 &= \sum_{j=1}^i d\phi_j + 2\pi \left(\sum_{j=1}^{i-1} \left(\sum_{k=1}^j df_k + \sum_{k=1}^{j-1} (j-k)d\dot{f}_k \Delta T \right) \Delta T + \frac{1}{2} \sum_{j=1}^{i-1} \sum_{k=1}^j \Delta \dot{f}_k \Delta T^2 \right) \\
 &= \sum_{j=1}^i d\phi_j + 2\pi \left(\sum_{j=1}^{i-1} (i-j)df_j \Delta T + \sum_{j=1}^{i-1} \sum_{k=1}^{j-1} (j-k)d\dot{f}_k \Delta T^2 + \frac{1}{2} \sum_{j=1}^{i-1} (i-j)\Delta \dot{f}_j \Delta T^2 \right) \\
 &= \sum_{j=1}^i d\phi_j + 2\pi \left(\sum_{j=1}^{i-1} (i-j)df_j \Delta T + \frac{1}{2} \sum_{j=1}^{i-1} ((i-j)(i-j-1)) + (i-j) d\dot{f}_j \Delta T^2 \right) \\
 &= \sum_{j=1}^i d\phi_j + 2\pi \left(\sum_{j=1}^{i-1} (i-j)df_j \Delta T + \frac{1}{2} \sum_{j=1}^{i-1} (i-j)^2 d\dot{f}_j \Delta T^2 \right). \tag{9.10}
 \end{aligned}$$

This result can be interpreted as the accumulated phase over a time $i\Delta T$ due to a random walk in the phase, frequency, and spin-down rate.

9.2 Random walk models: a simple treatment

We will now calculate the mismatch for a fully-coherent search given the random walk in phase, frequency, and spin-down rate defined in the previous section.

Let us begin by expanding the metric-mismatch summation from Eqn. (6.49). Recalling that Greek indices label the parameter components and Roman indices label the subdomain, then writing the summations explicitly, we have

$$\tilde{\mu} = g_{\alpha\beta ij} \Delta\lambda^{\alpha i} \Delta\lambda^{\beta j} \tag{9.11}$$

$$= \sum_{i=1}^{N_{\text{sd}}} \sum_{j=1}^{N_{\text{sd}}} g_{\alpha\beta ij} \Delta\lambda^{\alpha i} \Delta\lambda^{\beta j} \tag{9.12}$$

$$= \sum_{i=1}^{N_{\text{sd}}} g_{\alpha\beta ii} \Delta\lambda^{\alpha i} \Delta\lambda^{\beta i} + \sum_{i=1}^{N_{\text{sd}}} \sum_{j=1, j \neq i}^{N_{\text{sd}}} g_{\alpha\beta ij} \Delta\lambda^{\alpha i} \Delta\lambda^{\beta j}. \tag{9.13}$$

The summation has been intentionally split into terms for which the two subdomains are the same and those for which they are different. We calculated the metric when the reference time is at the beginning of each subdomain in Eqn. (6.57); by considering this metric for the two cases, we can define two distinct components as

$$g_{\alpha\beta ij} = \begin{cases} g_{\alpha\beta}^{\text{E}} & \text{if } i = j \\ g_{\alpha\beta}^{\text{NE}} & \text{if } i \neq j \end{cases}. \tag{9.14}$$

Finally, we can calculate the fully-coherent metric-mismatch from

$$\tilde{\mu} = \sum_{i=1}^{N_{\text{sd}}} g_{\alpha\beta}^{\text{E}} \Delta\lambda^{\alpha i} \Delta\lambda^{\beta i} + 2 \sum_{i=1}^{N_{\text{sd}}} \sum_{j=1}^{i-1} g_{\alpha\beta}^{\text{NE}} \Delta\lambda^{\alpha i} \Delta\lambda^{\beta j}. \quad (9.15)$$

9.2.1 Taking the expectation

In Eqn. (9.4), Eqn. (9.9), and Eqn. (9.10) we have written the parameter space offsets (which are to be used in calculating the mismatch) purely in terms of the random walk distributions $d\phi_i$, df_i , and $d\dot{f}_i$. We can calculate the mismatch exactly given a set of random walk jumps by inserting these into Eqn. (9.15). However, since we are dealing with statistical quantities, we can instead infer the behaviour of the mismatch under the random walk by taking an expectation.

Inserting Eqn. (9.4), Eqn. (9.9), Eqn. (9.10) in Eqn. (9.15) yields a number of terms with all the permutations of two terms from $[d\phi, df, d\dot{f}]$. Taking the expectation, all the cross-correlated terms, such as $d\phi_i d\dot{f}$, will have an expectation of zero since the steps of the random walk are independent. The only non-vanishing terms are given by

$$E[d\phi_i d\phi_j] = \delta_{ij} \sigma_\phi^2, \quad E[df_i df_j] = \delta_{ij} \sigma_f^2, \quad E[d\dot{f}_i d\dot{f}_j] = \delta_{ij} \sigma_{\dot{f}}^2, \quad (9.16)$$

After some simplification we find that the mismatch is given by

$$\begin{aligned} E[\tilde{\mu}] &= \frac{A_\phi}{6} \left(N_{\text{sd}} - \frac{1}{N_{\text{sd}}} \right) + \frac{\pi^2 A_f}{30} \left(4N_{\text{sd}}^3 + 5N_{\text{sd}}^2 + \frac{1}{N_{\text{sd}}} \right) \\ &\quad + \frac{\pi^2 A_{\dot{f}}}{3780} \left(66N_{\text{sd}}^5 - 21N_{\text{sd}}^3 + 105N_{\text{sd}}^2 + 217N_{\text{sd}} + 63 - \frac{94}{N_{\text{sd}}} \right), \end{aligned} \quad (9.17)$$

where

$$A_\phi = \sigma_\phi^2 \quad A_f = \sigma_f^2 \Delta T^2 \quad A_{\dot{f}} = \sigma_{\dot{f}}^2 \Delta T^4, \quad (9.18)$$

define three dimensionless ‘activity parameters’.

Recalling that $N_{\text{sd}} = T_{\text{obs}}/\Delta T$, Eqn. (9.17) makes predictions for the leading order scaling of the three random walks with the observation period

$$E[\tilde{\mu}]_{\text{PN}} \sim \sigma_\phi^2 \frac{T_{\text{obs}}}{\Delta T}, \quad E[\tilde{\mu}]_{\text{FN}} \sim \sigma_f^2 \frac{T_{\text{obs}}^3}{\Delta T}, \quad E[\tilde{\mu}]_{\text{SN}} \sim \sigma_{\dot{f}}^2 \frac{T_{\text{obs}}^5}{\Delta T}. \quad (9.19)$$

We note here the exact relation to the scaling of the variance of the root-mean phase residual as calculated by Cordes [43] and given in Eqn. (2.7). We will return to this later in Sec. 9.4.

9.2.2 Verifying the results

We can observe the leading order scaling of Eqn. (9.19) directly and verify the predictions made by Eqn. (9.17) by comparing with exact numerical results. That is, using the

LALSuite [106] signal injection and recovery tools developed in Section 8.3 of Chapter 8 we simulate signals undergoing a random walk and calculate the corresponding mismatch (no minimisation step is done here; this is discussed in the next section). In particular, we perform three Monte Carlo studies for a random walk in the phase, frequency, and spin-down rate and in each case compare the simulated results with the analytic prediction. The results are shown in Figure 9.2 and demonstrate good agreement between the simulation means and the prediction of Eqn. (9.17).

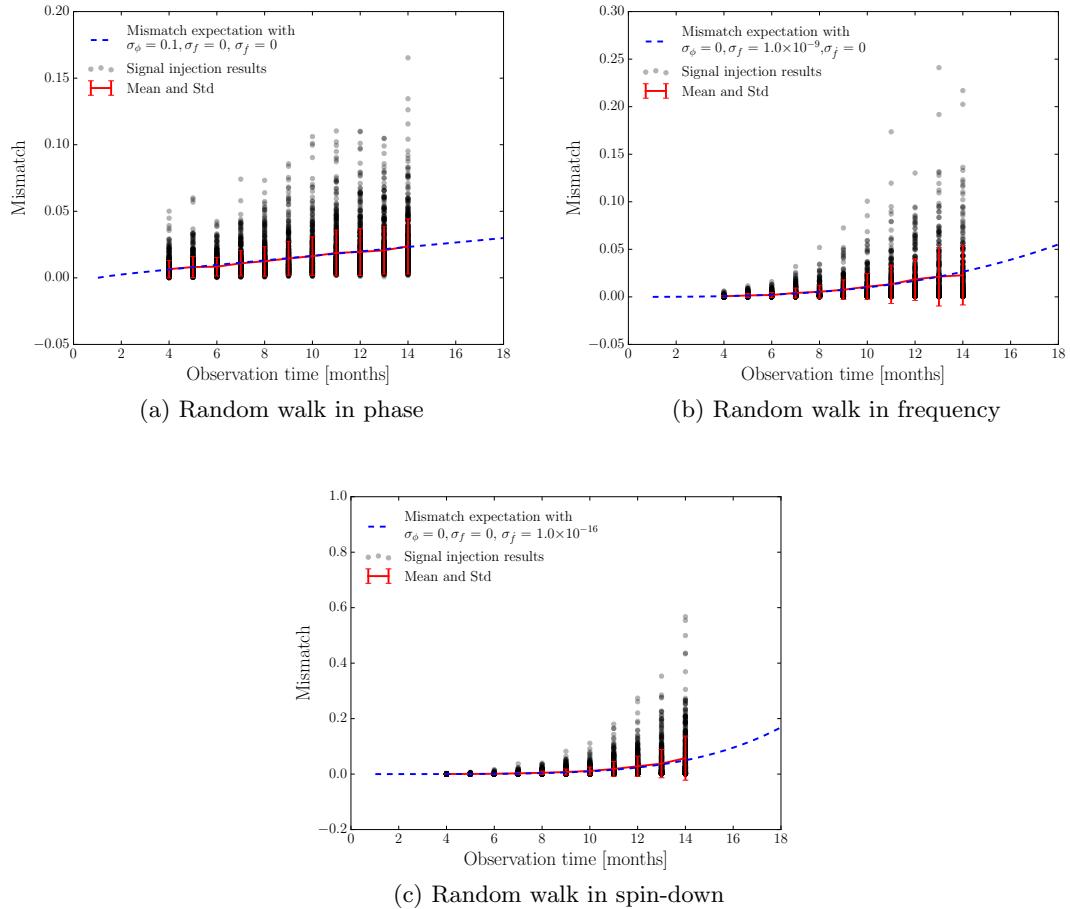


Figure 9.2: A comparison of Monte Carlo numerical simulated mismatch with the prediction of Eqn. (9.17) for a random walk in the phase, frequency, and spin-down rate.

9.3 Random walk models: minimising the mismatch

In Section 9.1 we have defined a random walk model for which we subsequently calculated the fully-coherent mismatch in Section 9.2. However, this is a special case in which the random walk for each parameter offset (the difference between the signal and the template) begins at the origin and then grows with time. It is the signal which undergoes a random walk, so in this case we have set the template to exactly match the signal at $t = 0$. However, one could imagine choosing a different template which would reduce

the overall mismatch; as such Eqn. (9.17) may overestimate the mismatch. The proper procedure is to minimise the mismatch with respect to the template parameters λ_t^α , which are implicitly in the calculation of Section 9.2, through the parameter space offset

$$\Delta\lambda^{\alpha i} = \lambda_s^{\alpha i} - \lambda_t^{\alpha i}, \quad (9.20)$$

as first defined in Section 6.4.

Ideally, we would like to repeat the calculation leading to Eqn. (9.17), minimising the mismatch with respect to the template parameters. However, this calculation is difficult and so we have not yet attempted it. A practical alternative method which we will use here is to begin with the random walk starting at the origin, as defined in Section 9.1, and then fit a polynomial of degree k , minimising the root-mean-square residual between the parameter space offset and the polynomial. We then subtract the best-fit polynomial from the parameter space offset. This leaves us with a residual parameter space offset for which we then compute the mismatch. We will verify that this captures the essential features of minimising the mismatch by comparing with numerical simulations in which the exact mismatch is minimised numerically.

In Appendix 9.B, we introduce the basic tools of least squares fitting and removing a polynomial of degree k to a generic random walk. In the following sections, we will calculate the minimised mismatch for random walks in the phase or frequency; we have not yet calculated the corresponding result for mixtures or random walks in the spin-down rate. In this process, we have a choice in the degree of polynomial to fit and remove. Since most searches minimise the mismatch with respect to the template frequency f_t and spin-down rate \dot{f}_t , this is equivalent to fitting and removing a $k = 2$ polynomial to the phase residual. In the following sections we will first calculate the analytic results and then verify that these agree with exact numerical results in Section 9.3.3.

9.3.1 Random walk in the phase

We begin with the simplest case of a random walk in phase, for which we have

$$\Delta\phi_i = \sum_{j=1}^i \mathcal{N}(0, \sigma_\phi^2). \quad (9.21)$$

Then, as shown in Eqn. (9.94) of Appendix 9.B, we have that

$$E[\Delta\phi_i \Delta\phi_j] = \sigma_\phi^2 \min(i, j). \quad (9.22)$$

We define the residual difference between the signal and template after fitting and removing the best-fit 2nd order polynomial, $\hat{y}_i^{(2)}$, as

$$\Delta^{(2)}\phi_i = \Delta\phi_i - \hat{y}_i^{(2)}. \quad (9.23)$$

Note that the superscript ‘(2)’ indicates the degree of polynomial and by $\Delta^{(2)}\phi_i$ we mean the residual difference between the signal and template after fitting and removing the polynomial.

We set the difference between the signal and template in all other parameters to zero such that the mismatch for a random walk in the residual phase is therefore

$$\tilde{\mu} = g_{00ij}\Delta^{(2)}\phi^i\Delta^{(2)}\phi^j \quad (9.24)$$

$$= \sum_{i=1}^{N_{\text{sd}}} g_{00}^E \Delta^{(2)}\phi_i \Delta^{(2)}\phi_i + 2 \sum_{i=1}^{N_{\text{sd}}} \sum_{j=1}^{i-1} g_{00}^{NE} \Delta^{(2)}\phi_i \Delta^{(2)}\phi_j. \quad (9.25)$$

To calculate the expectation of the mismatch, we need to evaluate the expectation of

$$\begin{aligned} \Delta^{(2)}\phi_i \Delta^{(2)}\phi_j &= \left(\Delta\phi_i - \sum_{k=1}^{N_{\text{sd}}} \mathcal{C}_{ik}^{(2)} \Delta\phi_k \right) \left(\Delta\phi_j - \sum_{l=1}^{N_{\text{sd}}} \mathcal{C}_{jl}^{(2)} \Delta\phi_l \right) \\ &= \Delta\phi_i \Delta\phi_j - \left(\sum_{k=1}^{N_{\text{sd}}} \mathcal{C}_{ik}^{(2)} \Delta\phi_j \Delta\phi_k + \sum_{l=1}^{N_{\text{sd}}} \mathcal{C}_{jl}^{(2)} \Delta\phi_i \Delta\phi_l \right) \\ &\quad + \sum_{k=1}^{N_{\text{sd}}} \sum_{l=1}^{N_{\text{sd}}} \mathcal{C}_{ik}^{(2)} \mathcal{C}_{jl}^{(2)} \Delta\phi_k \Delta\phi_l, \end{aligned} \quad (9.26)$$

where $\mathcal{C}_{ij}^{(2)}$ is defined in Eqn. (9.111) and Eqn. (9.110) of Appendix 9.B and we have replaced the Δx notation of the appendix with the time ΔT . Then taking the expectation

$$\begin{aligned} \mathbb{E} [\Delta^{(2)}\phi_i \Delta^{(2)}\phi_j] &= \sigma_\phi^2 \left(\min(i, j) - \left(\sum_{k=1}^{N_{\text{sd}}} \mathcal{C}_{ik}^{(2)} \min(j, k) + \sum_{l=1}^{N_{\text{sd}}} \mathcal{C}_{jl}^{(2)} \min(i, l) \right) \right. \\ &\quad \left. + \sum_{k=1}^{N_{\text{sd}}} \sum_{l=1}^{N_{\text{sd}}} \mathcal{C}_{ik}^{(2)} \mathcal{C}_{jl}^{(2)} \min(k, l) \right). \end{aligned} \quad (9.28)$$

Using symbolic mathematics packages we calculate an analytic expression which is a function of ΔT , i , j and N_{sd} . Inserting this into Eqn. (9.25) and simplifying we find that

$$E[\tilde{\mu}] = \sum_{i=1}^{N_{\text{sd}}} g_{00}^E E [\Delta^{(2)}\phi_i \Delta^{(2)}\phi_i] + 2 \sum_{i=1}^{N_{\text{sd}}} \sum_{j=1}^{i-1} g_{00}^{NE} E [\Delta^{(2)}\phi_i \Delta^{(2)}\phi_j] \quad (9.29)$$

$$= \frac{1}{70} \sigma_\phi^2 \left(3N_{\text{sd}} - \frac{27}{N_{\text{sd}}} \right). \quad (9.30)$$

This expression can be compared to Eqn. (9.17) ignoring the effect of the random walk in the frequency and spin-down rate. Notably, we retain the same leading order scaling of N_{sd} , but the overall coefficient is decreased; the same effect was found by Cordes [43] for the root-mean-square phase residual after removing a polynomial.

Rearranging the expression in the bracket demonstrates the mismatch is negative or zero for $1 \geq N_{\text{sd}} \geq 3$. This is a reflection of the minimum number of points needed in order to perform the quadratic fit. We will discuss this in more detail in Sec. 9.B.6 for the simpler case of fitting and removing a polynomial from a generic random walk.

9.3.2 Random walk in the frequency

For a random walk in the frequency we have an added complexity caused by the frequency offsets in the phase. For the frequency offset, we have

$$\Delta f_i = \sum_{j=1}^i \mathcal{N}(0, \sigma_f^2). \quad (9.31)$$

Recalling that we set the reference time at the beginning of each subdomain, then as in Section 9.1, the induced phase offset is

$$\Delta\phi_i = 2\pi \sum_{j=1}^{i-1} \Delta f_j \Delta T \quad (9.32)$$

$$= 2\pi \Delta T \sum_{j=1}^{i-1} \sum_{k=1}^j \mathcal{N}(0, \sigma_f^2) \quad (9.33)$$

$$= 2\pi \Delta T \sum_{j=1}^i (i-j) \mathcal{N}(0, \sigma_f^2). \quad (9.34)$$

Note that we do not include a random walk in the phase here.

Then we calculate the expected values of combinations of the parameter space offsets using Eqn. (9.94) from Appendix 9.B and the two summation identities Eqn. (9.76) and Eqn. (9.80) from Appendix 9.A. This gives

$$E[\Delta f_i \Delta f_j] = \sigma_f^2 \min(i, j), \quad (9.35)$$

$$E[\Delta\phi_i \Delta f_j] = 2\pi \Delta T \sigma_f^2 \sum_{k=1}^{\min(i,j)} (i-k), \quad (9.36)$$

$$E[\Delta\phi_i \Delta\phi_j] = (2\pi \Delta T)^2 \sigma_f^2 \sum_{k=1}^{\min(i,j)} (i-k)(j-k). \quad (9.37)$$

In Eqn. (9.23), we defined the residual difference between the signal and template phase after fitting and removing a second order polynomial. The second order polynomial was chosen to model the effect of minimising over the template frequency and frequency derivative. Let us now define

$$\Delta^{(1)} f_i = \Delta f_i - \hat{y}_i^{(1)}, \quad (9.38)$$

as the residual difference between the signal and template frequency after fitting and removing a first order polynomial. In this instance, the first order polynomial models the effect of minimising over the template frequency and frequency derivative.

To calculate the mismatch, we expand Eqn. (6.49) summing over the residual frequency offset $\Delta^{(1)} f_i$ (defined in Eqn. (9.38)) and the residual phase offset $\Delta^{(2)} \phi_i$ (given by

substituting Eqn. (9.34) into Eqn. (9.23)). This gives

$$\begin{aligned} E[\tilde{\mu}] = & \sum_{i=1}^{N_{\text{sd}}} \left(g_{00}^E E \left[\Delta^{(2)} \phi_i \Delta^{(2)} \phi_i \right] + 2g_{01}^E E \left[\Delta^{(2)} \phi_i \Delta^{(1)} f_i \right] + g_{11}^E E \left[\Delta^{(2)} f_i \Delta^{(1)} f_i \right] \right) \\ & + 2 \sum_{i=1}^{N_{\text{sd}}} \sum_{j=1}^{i-1} \left(g_{00}^{NE} E \left[\Delta^{(2)} \phi_i \Delta^{(2)} \phi_j \right] + g_{01}^{NE} E \left[\Delta^{(2)} \phi_j \Delta^{(2)} f_i \right] + \right. \\ & \quad \left. g_{10}^{NE} E \left[\Delta^{(2)} \phi_i \Delta^{(2)} f_j \right] + g_{11}^{NE} E \left[\Delta^{(1)} f_i \Delta^{(1)} f_j \right] \right). \end{aligned} \quad (9.39)$$

We calculate each of these expressions in a similar manner to Eqn. (9.28) replacing the relevant expectations with those given in Eqn. (9.35) to Eqn. (9.37). This yields an expected mismatch given by

$$E[\tilde{\mu}] = \frac{\pi^2}{630} \sigma_f^2 \Delta T^2 \left(N_{\text{sd}}^3 + 13N_{\text{sd}} + \frac{82}{N_{\text{sd}}} \right). \quad (9.40)$$

This can be compared with the frequency noise term alone in Eqn. (9.17). We note that the leading order power remains unchanged, but there is a reduction in the coefficient and a difference in the second highest power. The reduction in the coefficient is expected since we have minimised the mismatch; we do not have an intuitive explanation for the change in the second highest power.

9.3.3 Verifying the results

We now verify Eqn. (9.40) and Eqn. (9.30) by comparing with Monte Carlo simulations as we first did in Section 9.2.2. The numerical signals undergo a random walk as described in Section 9.2. However, when searching for the signals, we now search over a grid of points in f_t and \dot{f}_t . Then we select the grid point with the minimum mismatch; this minimises the mismatch over the frequency and spin-down. The results are plotted in Figure 9.3 and demonstrate reasonable agreement between the analytic prediction and the mean of the simulated mismatches.

9.4 Understanding the random walk model

In the previous two sections, we have used the random walk model defined in Section 9.1 which appears at first to be non-physical: the random walk ‘jumps’ every ΔT , but what physics determines this length? Furthermore, we assumed that the jumps between adjacent steps (see for example Eqn. (9.1)) were normally distributed with zero mean and variance σ_ϕ^2 , σ_f^2 , and $\sigma_{\dot{f}}^2$. But how physical is this assumption? In this section, we will answer both these questions and show that such a random walk is equivalent to the more natural ‘compound Poisson random walk process’ introduced in Section 2.3.1 which is used in the literature to model timing noise as a random walk. The methods

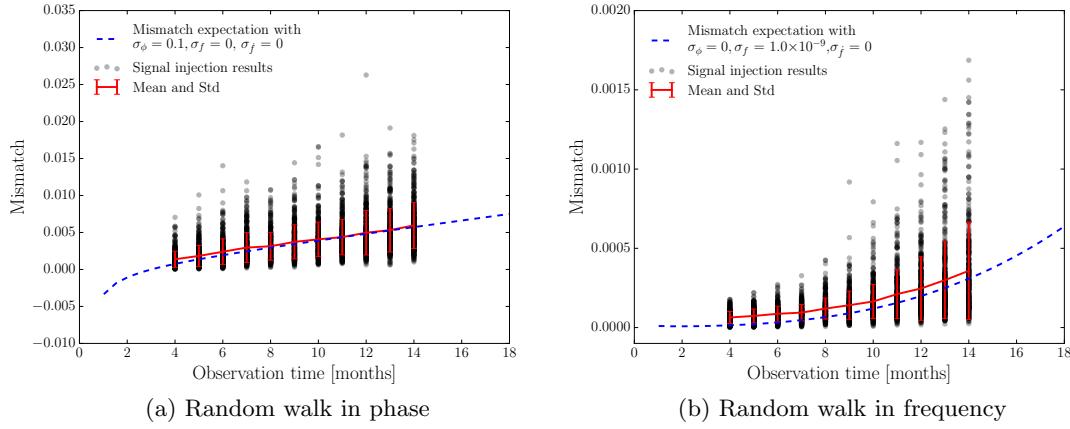


Figure 9.3: A comparison of the Monte Carlo numerical simulated mismatch with the predictions of Eqn. (9.30) and Eqn. (9.40); this differs from Figure 9.2 in that the numerical mismatch is minimised by selecting the smallest mismatch from a grid of points in f_t and \dot{f}_t .

discussed here parallel the work of Groth [73] and Cordes [43], but are described in detail as there are some subtleties to our analysis.

Consider a parameter $\Delta X(t) \in [\Delta\phi, \Delta\nu, \Delta\dot{\nu}]$ being the difference between the real noisy signal and the search template. $\Delta X(t)$ then encodes the deviations due to timing noise without concerning the secular spin-down. We will model timing noise by allowing $\Delta X(t)$ to undergo a *compound Poisson process*: a random walk where events are Poisson distributed in time occurring with a rate λ , but the size the events are drawn from a normal distribution with zero mean and a variance $\langle \delta X^2 \rangle$, or more formally

$$\Delta X(t) = \sum_{i=1}^{N(t)} \delta X_i, \quad (9.41)$$

where $\{N(t) : t \geq 0\}$ is a Poisson process with rate λ and $\{\delta X_i : i \geq 1\}$ are independent and identically distributed with $\delta X_i \sim N(0, \langle \delta X^2 \rangle)$. Note that $\langle \delta X^2 \rangle$ where $X \in [\phi, \nu, \dot{\nu}]$ is the variance of the random walk jumps.

If the event rate λ of the compound Poisson process is sufficiently large such that in a time ΔT the number of events is large, it can be shown that, due to the central limit theorem (see for example Weiss et al. [166]), $\Delta X_i(\Delta T)$ is normally distributed. The mean can be calculated using Wald's equation [161] which gives

$$\langle \Delta X(\Delta T) \rangle = 0, \quad (9.42)$$

while the variance can be calculated using the law of total variance [166] from which we find that

$$\langle \Delta X(\Delta T)^2 \rangle = \lambda \Delta T \langle \delta X^2 \rangle. \quad (9.43)$$

If we label the parameter offset measured for non-overlapping blocks of data ΔT as $\Delta X_i(\Delta T)$ then $\{\Delta X_1(\Delta T), \Delta X_2(\Delta T), \dots\}$ is a sequence of independent and identically distributed random variables with a mean and variance given by Eqn. (9.42) and Eqn. (9.43) respectively.

In the model defined in Section 9.1, we defined

$$dX_i = \Delta X_i - \Delta X_{i-1}, \quad (9.44)$$

as the difference between two adjacent subdomains of the piecewise Taylor expansion. Moreover, we assumed that $dX_i(\Delta T)$ was normally distributed with zero mean and a variance σ_X^2 . This assumption is valid if the underlying random walk process is a compound Poisson process. To see this, we note that if both ΔX_i and ΔX_{i-1} can be described by the compound Poisson process in Eqn. (9.41) and the assumption of a large number of events in each is valid, then

$$dX_i(\Delta T) = \Delta X_i \left(\frac{\Delta T}{2} \right) - \Delta X_{i-1} \left(\frac{\Delta T}{2} \right), \quad (9.45)$$

is the difference between normally distributed random variables. The factors of a half are due to splitting the period of $dX_i(\Delta T)$ into two equal durations. It can be shown (see for example Weisstein [168]) that the difference between two normal distributions, i.e. $dX_i(\Delta T)$, is also normally distributed with a mean which is the difference in the means of $\Delta X_i(\Delta T)$ and $\Delta X_{i-1}(\Delta T)$:

$$\langle dX(\Delta T) \rangle = 0, \quad (9.46)$$

and a variance which is the sum of the variance of $\Delta X_i(\Delta T)$ and $\Delta X_{i-1}(\Delta T)$:

$$\langle dX(\Delta T)^2 \rangle = 2\lambda \frac{\Delta T}{2} \langle \delta X^2 \rangle = \lambda \Delta T \langle \delta X^2 \rangle. \quad (9.47)$$

The metric-mismatch calculations derived in the previous two sections are written as a function of $\sigma_X^2 \in [\sigma_\phi^2, \sigma_f^2, \sigma_{\dot{f}}^2]$ which is the variance of $dX(\Delta T)$. We can therefore relate the two by

$$\lambda \Delta T \langle \delta X^2 \rangle = \sigma_X^2. \quad (9.48)$$

Rearranging this equation, we can equate $\lambda \langle \delta X^2 \rangle$ with a *GW random walk strength parameter* which we define as

$$S_{\text{GWPN}} := \lambda \langle \delta \phi^2 \rangle = \frac{\sigma_\phi^2}{\Delta T}, \quad (9.49)$$

$$S_{\text{GWFN}} := \lambda \langle \delta f^2 \rangle = \frac{\sigma_f^2}{\Delta T}, \quad (9.50)$$

$$S_{\text{GWSN}} := \lambda \langle \delta \dot{f}^2 \rangle = \frac{\sigma_{\dot{f}}^2}{\Delta T}. \quad (9.51)$$

We distinguish between the strength of gravitational wave noise (S_{GWPN} , S_{GWFN} and S_{GWSN}) and the usual strength of noise in the rotational parameters (S_{PN} , S_{FN} and $S_{\dot{\text{PN}}}$) defined in Eqn. (2.3) to Eqn. (2.5). To relate these, we need to define the mechanism of gravitational wave emission. In the canonical non-axisymmetric distortion mechanism, $f = 2\nu$. Therefore the underlying jumps in signals are related by

$$\delta\phi = 2\delta\varphi, \quad (9.52)$$

$$\delta f = 2\delta\nu, \quad (9.53)$$

$$\delta\dot{f} = 2\delta\dot{\nu}, \quad (9.54)$$

where φ , ν and $\dot{\nu}$ are the rotational phase, frequency, and spin-down rate. The variances of the GW parameters are therefore equal to the variances of the rotational parameters multiplied by a factor of 4. Using the definition the rotation noise strengths from Eqn. (2.3) to Eqn. (2.5), we see that

$$S_{\text{GWPN}} = 4S_{\text{PN}}, \quad (9.55)$$

$$S_{\text{GWFN}} = 4S_{\text{FN}}, \quad (9.56)$$

$$S_{\text{GWSN}} = 4S_{\dot{\text{SN}}}. \quad (9.57)$$

We can now write the expectation of the fully-coherent metric-mismatch due to a random walk in terms of these strength parameters. First, for a random walk in the phase given by Eqn. (9.30) we take just the leading order term and substitute in Eqn. (9.49). This results in

$$E[\tilde{\mu}] = \frac{3}{70} S_{\text{GWPN}} T_{\text{obs}}. \quad (9.58)$$

Second, for a random walk in the frequency, we take the leading order terms from Eqn. (9.40) and substitute using Eqn. (9.50). This gives

$$E[\tilde{\mu}] = \frac{\pi^2}{630} S_{\text{GWFN}} T_{\text{obs}}^3. \quad (9.59)$$

From this we see how, when the random walk is considered as a compound Poisson random walk, the mismatch does not depend on ΔT . This justifies the random walk treatment that was used in the earlier sections.

9.5 Application to the Crab pulsar

In Section 8.2, we introduced the Crab ephemeris. The regular and independent measurements of the frequency and spin-down rate in the Crab ephemeris provide a unique view of timing noise as a ‘jump’ in the phase, frequency, and spin-down rate each month. Specifically by a jump we mean the discontinuity at the interface between two months as illustrated in Figure 8.1. In this section, we will interpret data from the ephemeris in the context of a random walk model. Note that in this section, we discuss the rotational frequency and frequency derivatives, which we denote by ν and $\dot{\nu}$ which we will relate

to the gravitational wave frequency f using the non-axisymmetric emission model for which $f = 2\nu$.

9.5.1 Distribution of jumps in the Crab ephemeris

We begin with a purely empirical look at the distribution of jumps in the Crab ephemeris. We will use the jumps in frequency in the next section to interpret the Crab ephemeris in the context of a random walk.

From the Crab ephemeris, there are three distributions which we will calculate here: the jumps in frequency, spin-down rate, and phase. It is not meaningful to simply look at the difference in frequency (for example) between any two months without adjusting for the secular spin-down. But what we really want is the difference which occurs at the interface. To calculate this we define $\nu_i(t)$ as the frequency according to the i^{th} month as evaluated at time t . Then if $\Delta t_i = t_{i+1} - t_i$, the frequency jump between months is

$$d\nu_i = \nu_{i+1}(t_{i+1} - \Delta t_i/2) - \nu_i(t_i + \Delta t_i/2), \quad (9.60)$$

$$= \left[\nu_{i+1} - \frac{\Delta t_i}{2} \dot{\nu}_{i+1} + \left(\frac{\Delta t_i}{2} \right)^2 \ddot{\nu}_{i+1} \right] - \left[\nu_i + \frac{\Delta t_i}{2} \dot{\nu}_i + \left(\frac{\Delta t_i}{2} \right)^2 \ddot{\nu}_i \right]. \quad (9.61)$$

The Crab ephemeris does not include the second order spin-down, but since the first order spin-down is not constant we assume a constant average second order spin-down given by

$$\ddot{\nu}_{av} = \frac{1}{N} \sum_i \frac{\dot{\nu}_{i+1} - \dot{\nu}_i}{t_{i+1} - t_i}, \quad (9.62)$$

where N is the number of data points in the ephemeris file. Inserting this into the Taylor expansion, the second order terms cancel, leaving a frequency jump between months given by

$$d\nu_i = (\nu_{i+1} - \nu_i) - (\dot{\nu}_{i+1} + \dot{\nu}_i) \frac{\Delta t_i}{2}. \quad (9.63)$$

Note, this is not the difference in frequency between monthly updates, but the jump between months in frequency at the interface as illustrated in figure 8.1.

Calculating the result of Eqn. (9.63) for all the data points in the Crab ephemeris we plot the estimated probability density using the Gaussian kernel density estimate (KDE) method [92] in Figure 9.4A. Note that we have filtered out differences which occur over known glitches as described by Espinoza et al. [59]. This is because we are interested in the timing noise activity and not the effect of glitches themselves.

Moving now to the spin-down, we can calculate the jump between months in a similar way

$$d\dot{\nu}_i = \dot{\nu}_{i+1}(t_{i+1} - \Delta t_i/2) - \dot{\nu}_i(t_i + \Delta t_i/2) \quad (9.64)$$

$$= (\dot{\nu}_{i+1} - \dot{\nu}_i) - \ddot{\nu}_{av} \Delta t_i. \quad (9.65)$$

For the spin-down rate, we found a population of large negative jumps in the post-glitch periods and hence are probably not related to the timing noise activity that we are

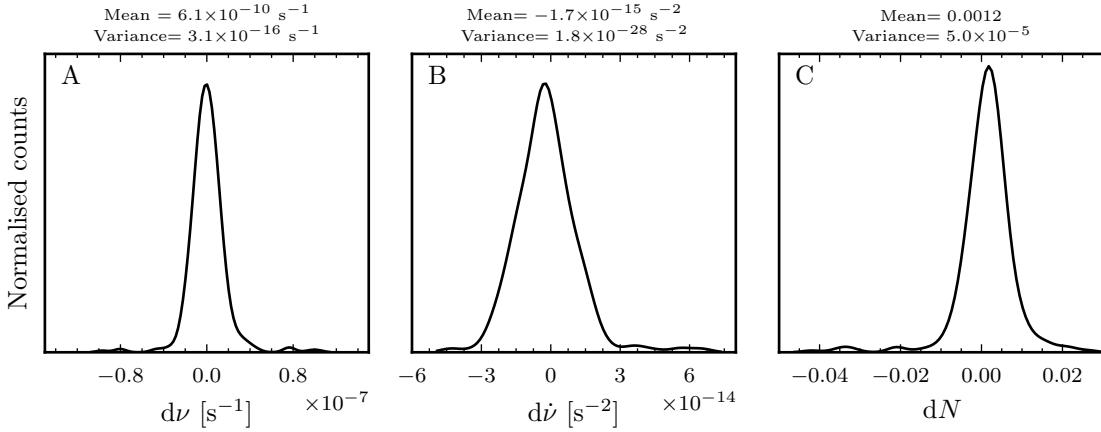


Figure 9.4: KDEs for the ‘jumps’ in frequency, spin-down rate, and residual number of rotations between adjacent per-month signals in the Crab ephemeris. Adjacent signals over glitches are filtered along with large anomalous values of dN once it was confirmed they occur within 200 days after a glitch.

interested in. As such in Figure 9.4B, we filter out these anomalous results to show the KDE for data points likely to be related to timing noise.

For the phase, we can use that each reference time given in the ephemeris coincides with a pulse. Therefore between two adjacent references times, the star has undergone an integer number of rotations. We can evaluate how well the ephemeris performs here by calculating the residual number of rotations between the timing model at a given step and the pulse arrival time at the next step. The data in the ephemeris file does not directly provide information on the phase evolution. It provides the independent phase evolution in each month with the phase at the reference time being zero. To calculate the full phase evolution between two reference times, we need to calculate the phase difference between each reference time and the interface time between them. Take this interface to be halfway between such that $t_{\text{mid}} = (t_i + t_{i+1})/2$. Then the total number of rotations between two reference times is

$$N = \frac{1}{2\pi} ((\phi_i(t_{\text{mid}}) - \phi_i(t_i)) - (\phi_{i+1}(t_{\text{mid}}) - \phi_{i+1}(t_{i+1}))). \quad (9.66)$$

We set the phase at the reference time to zero, so this leaves $N = \frac{1}{2\pi} (\phi_i(t_{\text{mid}}) - \phi_{i+1}(t_{\text{mid}}))$ where the terms are explicitly given by

$$\phi_i(t_{\text{mid}}) = 2\pi \left((t_{\text{mid}} - t_i)\nu_i + \frac{\dot{\nu}_i}{2!}(t_{\text{mid}} - t_i)^2 + \frac{\ddot{\nu}_{\text{av}}}{3!}(t_{\text{mid}} - t_i)^3 \right) \quad (9.67)$$

$$\phi_{i+1}(t_{\text{mid}}) = 2\pi \left((t_{\text{mid}} - t_{i+1})\nu_{i+1} + \frac{\dot{\nu}_{i+1}}{2!}(t_{\text{mid}} - t_{i+1})^2 + \frac{\ddot{\nu}_{\text{av}}}{3!}(t_{\text{mid}} - t_{i+1})^3 \right). \quad (9.68)$$

The total number of rotations N between months is a function of the length of a given month and the spin-down parameters. Calculating this allows us to check how well phase-connected lines of the ephemeris are. We quantify this by the residual number of

rotations, defined as

$$dN = N - \text{round}(N), \quad (9.69)$$

where by ‘round’ we indicate rounding to the nearest integer number of rotations.

In Figure 9.4C we plot the Gaussian KDE of ΔN having filtered against known glitch events. We found four jumps where $0.1 < \Delta N < 1.0$; again these were found to be occur within ~ 200 days of known glitches and so were removed to focus attention on the timing noise activity.

Figure 9.4 shows that the distribution of jumps in the frequency, spin-down rate, and residual number of rotations is centred on zero, as expected. The interesting part to note here is the size of the standard deviations. These can potentially be used to test timing noise models. In the next section, we interpret these in the context of a random walk.

9.5.2 Predicting the mismatch in the Crab

Studies of the Crab pulsar consistently find that the random walk noise is best fit by a frequency like noise process. In Table 9.1, we list the strengths reported from three works on the issue; this is by no means a complete literature review, but shows the typical values and spread.

	$S_{\text{FN}} \text{ Hz}^2/\text{s}$
Boynton et al. [33]	0.9×10^{-22}
Groth [73]	0.53×10^{-22}
Cordes [43]	0.66×10^{-22}

Table 9.1: Values for the strength of frequency noise in the Crab pulsar found in the literature: all three authors demonstrated that this strength was robust to changes in ΔT , the time which the data is divided into to calculate statistical quantities.

We can also derive our own estimate for the strength of frequency like timing noise in the Crab. To do this, we recall that from Eqn. (2.4) we have

$$S_{\text{FN}} = \lambda \langle \delta\nu^2 \rangle. \quad (9.70)$$

Then, as we did in Eqn. (9.50) for the gravitational wave frequency, we have that

$$S_{\text{FN}} = \frac{\langle d\nu(\Delta T)^2 \rangle}{\Delta T}. \quad (9.71)$$

Substituting $\langle d\nu(\Delta T)^2 \rangle = 3.1 \times 10^{-16} \text{ s}^{-1}$ (taken from Figure 9.4A) into Eqn. (9.71) with $\Delta T = 30$ days, this gives

$$S_{\text{FN}} = 1.20 \times 10^{-22} \text{ Hz}^2/\text{s}. \quad (9.72)$$

This strength is of the same order of magnitude as those listed in Table 9.1. We are not able to test if this strength is invariant to changes in ΔT because the Crab ephemeris is only calculated in $\Delta T = 1$ month long intervals; we therefore cannot confirm that the Crab undergoes frequency-like noise.

In Figure 8.5 we have already shown the dependence of the mismatch on the observation time for the Crab; fitting a power law to this we found that the mismatch scaled as

$$\langle \mu \rangle_{\text{fit}} \approx 1.5 \times 10^{-23} \left(\frac{T_{\text{obs}}}{1 \text{ sec}} \right)^{2.88}. \quad (9.73)$$

We can compare this directly with Eqn. (9.59). The power, being close to 3, suggests some similarity with the frequency like random walk in the mismatch. Now we can additionally predict this dependence given a value for the strength of frequency noise. This is done in Figure 9.5 for our estimated value of the strength parameter and using that $S_{\text{GWFN}} = 4S_{\text{FN}}$ (see Eqn. (9.56)). The prediction using the strength parameter

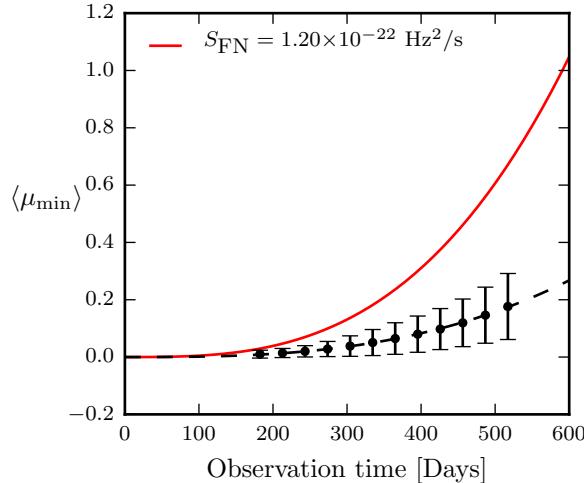


Figure 9.5: This figure, updated from Figure 8.5, shows the dependence of the minimum mismatch on the observation time for the Crab, along with the prediction of Eqn. (9.59) for the our estimate of the frequency noise strength parameter calculated in Eqn. (9.72).

calculated from the Crab ephemeris, as given in Eqn. (9.72), overestimates the dependence of the mismatch calculated exactly from the data. It is unclear at the time what causes this, but we have nevertheless demonstrated the essential elements of making such predictions which may be useful in the future, if there is a clear case where the timing noise is well modelled by a random walk.

9.6 Conclusion

In this Chapter, we have calculated the expectation of the fully-coherent mismatch when searching for a GW signal which undergoes a random walk in one of the phase, frequency,

or spin-down rate. We did this first for a system in which the difference between the signal and template was initially zero and then grew as a random walk. Since this is not a minimised result, we then demonstrated how to minimise the mismatch with respect to the template parameters f_t and \dot{f}_t . The formulae derived in this section have been verified against Monte-Carlo type simulations of the exact mismatch.

In Section 9.4, we related the random walk model to a more physical compound Poisson random walk. Using this description, we formulated the two key results of this chapter: Eqn. (9.58) and Eqn. (9.59). These are the expectation of the minimum metric-mismatch due to random walk in the phase and frequency, respectively, as a function of the strength of phase and frequency noise.

Following this, we developed our understanding of random walk models in the context of the data from the Crab ephemeris. We showed that the frequency noise strength can be measured from this data and has a value consistent with other results in the literature. We then demonstrated the use of this derived prediction of the strength of frequency noise and Eqn. (9.59) to predict the results found in Chapter 8 for the dependence of the mismatch on the observation time. The prediction is not perfect, but does provide a rough order of magnitude figure of the expected mismatch.

Appendix 9.A Summation identities

In this appendix, we derive two useful summation identities used in Sec. 9.2. First, we have

$$\sum_{b=1}^c \sum_{a=1}^b X_a = (X_1) + (X_1 + X_2) + \dots + (X_1 + X_2 + \dots + X_{c-1} + X_c) \quad (9.74)$$

$$= cX_1 + (c-1)X_2 + \dots + 2X_{c-1} + X_c \quad (9.75)$$

$$= \sum_{b=1}^c (c+1-b)X_b, \quad (9.76)$$

which is used in deriving Eqn. (9.9). Second, we have

$$\begin{aligned} \sum_{j=1}^{i-1} \sum_{k=1}^{j-1} (j-k) X_k &= [0] + [X_1] + [2X_1 + X_2] + [3X_1 + 2X_2 + X_3] + \dots \\ &\quad + [(i-2)X_1 + (i-3)X_2 + (i-4)X_3 + \dots \\ &\quad + 3X_{i-4} + 2X_{i-3} + X_{i-2}] \end{aligned} \tag{9.77}$$

$$\begin{aligned} &= (1+2+3+\dots+(i-4)+(i-3)+(i-2))X_1 \\ &\quad + (1+2+3+\dots+(i-4)+(i-3))X_2 \\ &\quad + (1+2+3+\dots+(i-4))X_3 + \dots \\ &\quad + (1+2+3)X_{i-4} + (1+2)X_{i-3} + X_{i-2} \end{aligned} \tag{9.78}$$

$$= \sum_{k=1}^{i-2} kX_1 + \sum_{k=1}^{i-3} kX_2 + \dots + \sum_{k=1}^2 kX_{i-3} + \sum_{k=1}^1 kX_{i-2} \tag{9.79}$$

$$= \sum_{j=1}^{i-2} \left(\sum_{k=1}^{i-1-j} k \right) X_j = \frac{1}{2} \sum_{j=1}^{i-2} (i-j)(i-j-1)X_j, \tag{9.80}$$

which is used in deriving Eqn. (9.10).

Appendix 9.B Least-squares minimisation of a random walk

In this appendix, we will describe the process of fitting and removing a polynomial from N data points (x_i, y_i) which undergoes a random walk. The polynomial will be fitted using a least squares minimisation. The x_i are the independent points at which y_i (which undergoes a random walk) is measured. We begin by defining the least-squares fitting method then go on to calculate the residual for several different degrees of polynomial. This introduces the method in a generic setting which is then applied in Section 9.3 to calculate the minimised mismatch for a GW signal undergoing a random walk.

9.B.1 Least squares fitting of a polynomial

Given N data points x_i, y_i , we define the residual from a least-squares polynomial fit of order k , as

$$r_i^{(k)} = y_i - y_i^{(k)}, \tag{9.81}$$

where

$$y_i^{(k)} = a_0 + a_1 x_i + a_2 x_i^2 + \dots + a_k x_i^k, \tag{9.82}$$

is a polynomial of degree k .

Then the residual which we want to minimise is

$$R^2 = \sum_{i=1}^N \left(r_i^{(k)} \right)^2 = \sum_{i=1}^N \left(y_i - \left(a_0 + a_1 x_i + a_2 x_i^2 + \dots + a_k x_i^k \right) \right)^2. \tag{9.83}$$

Partial differentiation with respect to the parameters a_i , yields k simultaneous equations. Writing these as a matrix and then solving for the best fit, $\hat{y}^{(k)}_i$, it can be shown (see for example Weisstein [167]) that

$$\hat{y}_i^{(k)} = X (X^T X)^{-1} X^T y_i \quad \text{where} \quad X = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^k \\ 1 & x_2 & x_2^2 & \dots & x_2^k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^k \end{bmatrix} \quad (9.84)$$

Here X is an example of a *Vandermonde* matrix in which the terms follow a geometric progression. It is useful to note that

$$XX^T = \begin{bmatrix} N & \sum_{i=1}^N x_i & \dots & \sum_{i=1}^N x_i^k \\ \sum_{i=1}^N x_i & \sum_{i=1}^N x_i^2 & \dots & \sum_{i=1}^N x_i^{k+1} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^N x_i^k & \sum_{i=1}^N x_i^{k+1} & \dots & \sum_{i=1}^N x_i^{2k} \end{bmatrix}. \quad (9.85)$$

Provided that the x_i are suitably defined, then an analytic fit can be found for any k , the difficulty lies in inverting the matrix.

9.B.2 Least squares fitting a polynomial to a random walk

We now take the x_i, y_i to be a Gaussian random walk beginning at the origin. To define this, let $dy_i \sim N(0, \sigma^2)$ be independent and identically distributed random variables for which their sum generates the random walk:

$$y_i = \sum_{j=1}^i dy_j. \quad (9.86)$$

We also set each random walk event to occur according to $x_i = i\Delta x$. Then the residual after fitting and removing a k^{th} order polynomial to the random walk y_i , is

$$r_i^{(k)} = y_i - \hat{y}_i^{(k)} = y_i - X (X^T X)^{-1} X^T y_i. \quad (9.87)$$

This suggests the residual will be similar to the random walk, but modified by the least squares fitting. To illustrate this, in Figure 9.6 we plot a simulated random walk along with several fits.

9.B.3 Zeroth order fitting

We begin with the case of $k = 0$ in which $X^T = [1, 1, \dots, 1]$ such that

$$X (X^T X)^{-1} X^T = \frac{1}{N} J_N \quad (9.88)$$

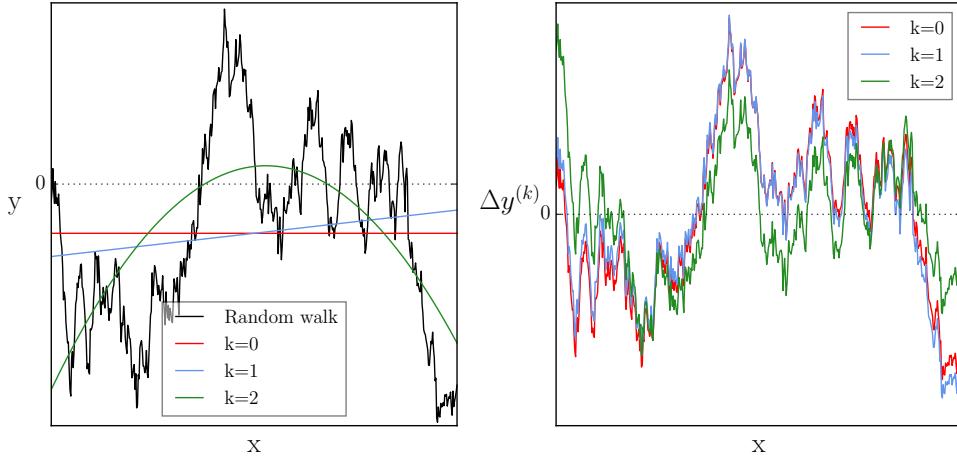


Figure 9.6: Example of a random walk on the left along with three polynomial fits of varying order. On the right is the corresponding residual after subtracting these fits. A dotted line marks the origin in both plots.

where J_N is the $N \times N$ matrix of ones. Inserting this into Eqn. (9.87), the residual from a zeroth order fit is given by

$$r_i^{(0)} = y_i - \frac{1}{N} \sum_{j=1}^N y_j. \quad (9.89)$$

The zeroth order residual can be interpreted as the removing the average value $\langle y_i \rangle$ from the random walk: this was illustrated in Figure 9.6.

We can now take expectations to understand the behaviour of the residual when compared to the original definition of the random walk in Eqn. (9.86). For example, consider the mean square translation distance from the origin of a random walk after i steps. For a normal random walk, this has the well known result

$$E[y_i^2] = i\sigma^2. \quad (9.90)$$

We can calculate the corresponding quantity of the $k = 0$ residual by first noting that

$$E[y_i y_j] = E \left[\sum_{k=1}^i dy_k \sum_{l=1}^j dy_l \right] \quad (9.91)$$

$$= \sum_{k=1}^i \sum_{l=1}^j E[dy_k dy_l] \quad (9.92)$$

$$= \sum_{k=1}^i \sum_{l=1}^j \delta_{kl} \sigma^2 \quad (9.93)$$

$$= \sigma^2 \min(i, j), \quad (9.94)$$

where δ_{kl} is the Kronecker delta. Then we have

$$\left(r_i^{(0)}\right)^2 = y_i^2 - \frac{2}{N} \sum_{k=1}^N y_i y_k + N^{-2} \sum_{k=1}^N \sum_{l=1}^N y_k y_l \quad (9.95)$$

$$= y_i^2 - 2N^{-1} \left(\sum_{k=1}^i y_i y_k + \sum_{k=i+1}^N y_i y_k \right) + N^{-2} \sum_{k=1}^N \left(\sum_{l=1}^k y_k y_l + \sum_{l=k+1}^N y_k y_l \right). \quad (9.96)$$

Taking the expectation we have

$$E \left[\left(r_i^{(0)} \right)^2 \right] = \sigma^2 \left(i - \frac{2}{N} \left(\sum_{k=1}^i k + \sum_{k=i+1}^N k \right) + \frac{1}{N^2} \sum_{k=1}^N \left(\sum_{l=1}^k l + \sum_{l=k+1}^N k \right) \right) \quad (9.97)$$

$$= \sigma^2 \left(\frac{N}{3} - i + \frac{1}{2} + \frac{i^2}{N} - \frac{i}{N} + \frac{1}{6N} \right). \quad (9.98)$$

This result can be compared with Eqn. (9.90), the expectation of the squared value for a random walk. In contrast, the expectation after i steps for the residual random walk depends on the length of data N that was fitted. It can be shown the expectation has a minimum at $i = N/2$.

To further understand the difference between the random walk and the residual random walk, let us consider the sum of squares after N steps for the random walk

$$E \left[\sum_{i=1}^N y_i^2 \right] = \sum_{i=1}^N i \sigma^2 = \frac{1}{2} (N^2 + N) \sigma^2. \quad (9.99)$$

On the other hand, the sum of squares for the residual random walk is given by

$$E \left[\sum_{i=1}^N \left(r_i^{(0)} \right)^2 \right] = \frac{1}{6} (N^2 - 1) \sigma^2. \quad (9.100)$$

Comparing equations (9.99) and (9.100) we note that, for the leading order term, the coefficient is reduced, but the power remains the same.

9.B.4 First order fitting

We now consider a first order fitting for which

$$\hat{y}_i^{(1)} = X (X^T X)^{-1} X^T y_i \quad \text{with} \quad X = \begin{bmatrix} 1 & \Delta x \\ 1 & 2\Delta x \\ \vdots & \vdots \\ 1 & N\Delta x \end{bmatrix}. \quad (9.101)$$

Inserting the definitions of x_i we can write

$$(X^T X)^{-1} = \frac{1}{N(N-1)} \begin{bmatrix} 4N+2 & -\frac{6}{\Delta x} \\ -\frac{6}{\Delta x} & \frac{12}{\Delta x^2(N+1)} \end{bmatrix} = C^{(1)}. \quad (9.102)$$

For convenience we have defined a symmetric matrix $C^{(1)}$. We then proceed to define another matrix

$$\mathcal{C}_{ij}^{(1)} := XC^{(1)}X^T \quad (9.103)$$

$$= \begin{bmatrix} 1 & \Delta x \\ 1 & 2\Delta x \\ \vdots & \vdots \\ 1 & N\Delta x \end{bmatrix} \begin{bmatrix} C_{11}^{(1)} & C_{12}^{(1)} \\ C_{21}^{(1)} & C_{22}^{(1)} \end{bmatrix} \begin{bmatrix} 1 & 1 & \dots & 1 \\ \Delta x & 2\Delta x & \dots & N\Delta x \end{bmatrix} \quad (9.104)$$

$$= C_{11}^{(1)} J_N + C_{12}^{(1)} \Delta x \begin{bmatrix} 2 & 3 & \dots & N+1 \\ 3 & 4 & \dots & \vdots \\ \vdots & & & \\ N+1 & \dots & \dots & 2N \end{bmatrix} + C_{22}^{(1)} \Delta x^2 \begin{bmatrix} 1 & 2 & \dots & N \\ 2 & 4 & \dots & \vdots \\ \vdots & & & \\ N & \dots & \dots & N^2 \end{bmatrix} \quad (9.105)$$

We can write $r_i^{(1)}$ as a summation by inferring the dependence of the i^{th} row of each matrix on the j^{th} column

$$r_i^{(1)} = y_i - \sum_{j=1}^N \mathcal{C}_{ij}^{(1)} y_j \quad \text{where} \quad \mathcal{C}_{ij}^{(1)} = C_{11}^{(1)} + C_{12}^{(1)} \Delta x(i+j) + C_{22}^{(1)} \Delta x^2 ij \quad (9.106)$$

We have now defined the first order residual. To understand that fitting an removing a first order polynomial has, we compute the expectation of the square for the i^{th} term

$$\begin{aligned} E \left[\left(r_i^{(1)} \right)^2 \right] &= \frac{1}{15N(N^2-1)} (2N^4 - 18N^3i + 9N^3 + 78N^2i^2 - 78N^2i \\ &\quad + 14N^2 - 120Ni^3 + 180Ni^2 - 78Ni \\ &\quad + 9N + 60i^4 - 120i^3 + 78i^2 - 18i + 2), \end{aligned} \quad (9.107)$$

which can be compared with the classic result for a random walk given in Eqn. (9.90). Alternatively, comparing with Eqn. (9.99), the expected sum of squares for the residual random walk is

$$E \left[\sum_{i=1}^N \left(r_i^{(1)} \right)^2 \right] = \frac{1}{15} (N^2 - 4) \sigma^2. \quad (9.108)$$

As in the zeroth order fit, the power of the leading order term remains the same, but the coefficient decreases.

9.B.5 Second order fitting

For the residual left after removing a quadratic, the argument proceeds in much the same way with

$$r_i^{(2)} = y_i - \sum_{j=1}^N C_{ij}^{(2)} y_j, \quad (9.109)$$

where

$$\begin{aligned} C_{ij}^{(2)} = & C_{11}^{(2)} + C_{22}^{(2)} \Delta x^2 i j + C_{33}^{(2)} \Delta x^4 i^2 j^2 + C_{12}^{(2)} \Delta x (i + j) \\ & + C_{13}^{(2)} \Delta x (i^2 + j^2) + C_{23}^{(2)} \Delta x^3 (i j^2 + i^2 j), \end{aligned} \quad (9.110)$$

and

$$C^{(2)} = \frac{1}{N(N-1)(N-2)} \begin{bmatrix} 9N^2 + 9N + 6 & -\frac{1}{\Delta x} (36N + 18) & \frac{30}{\Delta x^2} \\ -\frac{1}{\Delta x} (36N + 18) & \frac{12(2N+1)(8N+11)}{\Delta x^2(N+1)(N+2)} & -\frac{180}{\Delta x^3(N+2)} \\ \frac{30}{\Delta x^2} & -\frac{180}{\Delta x^3(N+2)} & \frac{180}{\Delta x^4(N+1)(N+2)} \end{bmatrix}. \quad (9.111)$$

The expression for the expected square value is too long to write out in full, but the expected sum of squares for the residual random walk is

$$E \left[\sum_{i=1}^N (r_i^{(2)})^2 \right] = \frac{1}{70} (3N^2 - 27) \sigma^2, \quad (9.112)$$

for which as in the case of the zeroth order and first order residuals, the power of the leading order term remains unchanged, but the coefficient decreases.

9.B.6 Conclusions

We now have a method to calculate statistical quantities from the residual left over after subtracting a k^{th} order polynomial from a random walk. Considering the sum of squares for a random walk and the residuals in equations (9.99), (9.100), (9.108), and (9.112) we find that the leading order term retains the same power of N with increasing k but the coefficient of this power gets smaller. This reflects the improved fitting with the polynomial degrees. We also note that with each increase in the order of fit we get a limit on N for which the sum of squares is positive. For zeroth order fitting this is $N > 1$, for first order $N > 2$ and for second order $N > 3$. This is because in order to perform a least squares fit, we need at least $k + 1$ points to fit.

Chapter 10

Conclusion and outlook

Timing variations in pulsars have long been used as a way to infer the elemental properties of neutron stars. However, while in isolated cases models for glitches and timing noise have successfully explained the observations, we are some way from having a complete and universal understanding of neutron stars. For timing noise, this may be due to the varied ways in which it manifests, but in the comprehensive review by Hobbs et al. [81] the authors suggest that over sufficiently long timescales (which may be much longer than we are able to observe) timing residuals show quasi-periodic features. Moreover, some pulsars (see Lyne et al. [111] for examples) have pronounced periodic modulations in not only their timing residuals, but also the shape of pulsations, PSR B1828-11 is the best example of this. With this in mind, in this thesis we investigated two mechanisms of strictly periodic variations in detail: magnetospheric switching and precession.

We began in Chapter 3 by studying the action of the electromagnetic torque on precessing pulsars using numerical solutions to the Euler rigid-body equations coupled to the Deutsch [54] electromagnetic torque. This allowed us to investigate the role of the anomalous component of the torque and conclude that, for realistic neutron stars, this can safely be ignored. We cleared up some confusion in the literature, by showing that the ‘persistent precession’ solutions found by Melatos [122] were in fact not precessing, but aligned with the principal axes of the effective body frame arising from the inclusion of the anomalous torque.

In order to test models, they need to be predictive, in the sense that they model the features of a neutron star observed by pulsar astronomers. To address this in the context of precession, in Chapter 4 we developed numerical solutions which, together with solving the Euler rigid-body equations, solve for the Euler rotation angles which transform from the rotating frame to the inertial frame. This allowed us to directly model the phase residual, spin-down rate, and pulse profile of a precessing pulsar. We compared the numerical model against analytic solutions from the literature and against our derivation of the spin-down rate which we used later in Chapter 5. We finished this chapter by discussing some preliminary results from a hybrid model which couples precession to magnetospheric torque switching. In the future, it would be interesting to explore this further by understanding better how reconfiguring the magnetosphere will effect the

Deutsch [54] torque and then testing different mechanisms to bias, in a probabilistic way, the switching using the precessional cycle. The numerical model is perfectly suited to this task as it captures the complicated feedback between the torque switching and precession.

PSR B1828-11 stands out in the literature for the strong ~ 500 day periodic modulations observed in its timing properties. As a result, it has been used as evidence for both precession [156] and periodic magnetospheric torque switching [111]. Jones [89] and Cordes [45] suggested that these models may not be mutually exclusive, which prompted our hybrid model discussed in Chapter 4. Neglecting these hybrid models, precession and magnetospheric switching are mutually exclusive, but it is important to know which is favoured since both have important implications for neutron star physics. To decide this, in Chapter 5, we applied a Bayesian model comparison. We found an odds-ratio of $10^{2.7 \pm 0.5}$ in favour of the precession model, a key result of this thesis which we published in Ashton et al. [25]. This does not rule out the switching interpretation entirely as we have not tested an exhaustive set of models, but it does provide a quantitative framework to evaluate models. In this chapter we focus primarily on the methods used to ensure that we make an unbiased comparison, although there is some development for the modulations of the beam-width due to precession. In the future, we intend to use these tools to test further modifications of the models. Furthermore, we would like to combine the data from other pulsars with long period modulations, such as PSR 0919+06 [134], and repeat the model comparison to see how this changes the odds-ratio.

Detecting gravitational waves from an isolated neutron star would provide a unique opportunity to learn about them, especially if we additionally observed the object from its electromagnetic output. However, signals which are subject to timing variations may be difficult to detect if, as in the case of many current gravitational wave searches, we use matched filtering templates which do not include timing variations. This issue has not been properly tackled in the literature and so in the last chapters of this thesis we begin the process of quantifying the risks posed by timing variations to efforts to detect gravitational waves from neutron stars.

In Chapter 6 we set out the tools which we will use to calculate the mismatch, defined to be the loss of signal to noise ratio. In doing so we defined a new approach, the generalised metric-mismatch, which can calculate the mismatch for any arbitrary signal by approximating it as a piecewise Taylor expansion.

Searches for gravitational waves from known pulsars can handle glitches in the signal since they can see that they have occurred. In Chapter 7 we showed that for blind searches, where we have no electromagnetic signal, glitches pose a substantial risk, especially since many searches target young rapidly spinning-down pulsars which we found to be likely to have larger and more frequent glitches. Many of these searches use an initial semi-coherent stage and then follow-up candidates with fully-coherent searches. We showed how a signal can be identified as a candidate in the semi-coherent stage and then subsequently lost in the follow-up. In the future, we would like to develop search methods which are robust to glitches. These could then be applied to past data to ensure that we have not already missed a continuous gravitational wave signal.

Quantifying the effect of timing noise on gravitational wave searches is a more difficult task due to the fact that we have no universal empirical description of what constitutes timing noise; therefore, it is difficult to predict what timing noise may exist in gravitational wave signals from isolated neutron stars. One way to approach this, discussed in Chapter 8, is to use data on the frequency and spin-down rate evolution of the Crab pulsar to generate ‘noisy gravitational wave signals’. We searched for these using standard matched filtering tools and calculated the minimum mismatch having searched in a narrow band of frequency and spin-down rate. The conventional wisdom for continuous wave searches contends that, since the SNR scales with the square-root of the observation time, provided we look for a sufficiently long time the signal will eventually become detectable. However, by investigating how the minimum mismatch scaled with the observation time for noisy signals, we found that $\langle \tilde{\mu}_{\min} \rangle \propto T_{\text{obs}}^{2.88}$. This result, published in Ashton et al. [24], means that for noisy signals the conventional wisdom does not hold. There is in fact an observation time after which the SNR decreases with observation time; for the Crab pulsar this was found to be ≈ 600 days.

In Chapter 9 we approached the issue of timing noise in continuous gravitational waves from a different angle. Namely, by modelling timing noise as a random walk in the phase, frequency, or spin-down rate: a phenomenological description dating back to Boynton et al. [33]. Recently, this model has been disfavoured [81] as a substantive explanation of timing noise, but it nevertheless provides a simple empirical description which is consistent with the timing residuals of many pulsars. Moreover, while we apply it to isolated neutron stars, it can also be applied to low-mass X-ray binary systems where spin-wandering due to fluctuations in the torque can be modelled by a random walk. In this chapter, we calculated the mismatch due to random walks and showed how to include the minimisation step modelling a narrow-band search over the template frequency and spin-down rate. We go on to use the Crab ephemeris to predict the strength of frequency noise in the Crab pulsar and hence the dependence of the mismatch on observation time found in Chapter 8. This will be useful in developing models that we would like to use in the future to estimate the levels of mismatch for other searches which may be affected by random walk models of timing noise.

In this thesis, we hope to have developed the understanding of timing noise in neutron stars. Most significantly, by providing a framework with which to debate the merits of models explaining periodic signals and quantify their comparison. We also hope that by understanding the role of timing variations in continuous gravitational waves, we help to guide efforts to detect these elusive signals in the future.

Bibliography

- [1] Aasi, J., Abadie, J., Abbott, B. P., Abbott, R., Abbott, T., Abernathy, M. R., Accadia, T., Acernese, F., Adams, C., Adams, T., et al. (2013a). Directed search for continuous gravitational waves from the Galactic center. *Physical Review D*, 88(10):102002.
- [2] Aasi, J., Abadie, J., Abbott, B. P., Abbott, R., Abbott, T., Abernathy, M. R., Accadia, T., Acernese, F., Adams, C., Adams, T., et al. (2014). Application of a Hough search for continuous gravitational waves on data from the fifth LIGO science run. *Classical and Quantum Gravity*, 31(8):085014.
- [3] Aasi, J., Abadie, J., Abbott, B. P., Abbott, R., Abbott, T. D., Abernathy, M., Accadia, T., Acernese, F., Adams, C., Adams, T., et al. (2013b). Einstein@Home all-sky search for periodic gravitational waves in LIGO S5 data. *Physical Review D*, 87(4):042001.
- [4] Aasi, J., Abbott, B. P., Abbott, R., Abbott, T., Abernathy, M. R., Acernese, F., Ackley, K., Adams, C., Adams, T., Addesso, P., et al. (2015a). Directed search for gravitational waves from Scorpius X-1 with initial LIGO data. *Physical Review D*, 91(6):062008.
- [5] Aasi, J., Abbott, B. P., Abbott, R., Abbott, T., Abernathy, M. R., Acernese, F., Ackley, K., Adams, C., Adams, T., Addesso, P., et al. (2015b). Narrow-band search of continuous gravitational-wave signals from Crab and Vela pulsars in Virgo VSR4 data. *Physical Review D*, 91(2):022004.
- [6] Aasi, J., Abbott, B. P., Abbott, R., Abbott, T., Abernathy, M. R., Acernese, F., Ackley, K., Adams, C., Adams, T., Addesso, P., et al. (2015c). Searches for continuous gravitational waves from nine young supernova remnants. *The Astrophysical Journal*, 813(1):39.
- [7] Aasi, J., Abbott, B. P., Abbott, R., Abbott, T., Abernathy, M. R., Acernese, F., Ackley, K., Adams, C., Adams, T., Addesso, P., et al. (2016). First low frequency all-sky search for continuous gravitational wave signals. *Physical Review D*, 93(4):042007.
- [8] Abadie, J., Abbott, B., Abbott, R., Abbott, T., Abernathy, M., Accadia, T., Acernese, F., Adams, C., Adhikari, R., Affeldt, C., et al. (2012). All-sky search for periodic gravitational waves in the full S5 LIGO data. *Physical Review D*, 85(2):022001.
- [9] Abadie, J., Abbott, B. P., Abbott, R., Abernathy, M., Accadia, T., Acernese, F., Adams, C., Adhikari, R., Affeldt, C., Allen, B., et al. (2011). Beating the spin-down

- limit on gravitational wave emission from the Vela pulsar. *The Astrophysical Journal*, 737(2):93.
- [10] Abbott, B., Abbott, R., Adhikari, R., Agresti, J., Ajith, P., Allen, B., Amin, R., Anderson, S. B., Anderson, W., Arain, M., et al. (2007). Searches for periodic gravitational waves from unknown isolated sources and Scorpius X-1: Results from the second LIGO science run. *Physical Review D*, 76(8):082001.
- [11] Abbott, B., Abbott, R., Adhikari, R., Ajith, P., Allen, B., Allen, G., Amin, R., Anderson, S. B., Anderson, W. G., Arain, M. A., et al. (2008). Beating the spin-down limit on gravitational wave emission from the Crab pulsar. *The Astrophysical Journal Letters*, 683(1):L45.
- [12] Abbott, B. P., Abbott, R., Abbott, T., Abernathy, M., Acernese, F., Ackley, K., Adams, C., Adams, T., Addesso, P., Adhikari, R. X., et al. (2016). Observation of gravitational waves from a binary black hole merger. *Physical review letters*, 116(6):061102.
- [13] Abbott, B. P., Abbott, R., Acernese, F., Adhikari, R., Ajith, P., Allen, B., Allen, G., Alshourbagy, M., Amin, R., Anderson, S., et al. (2010). Searches for gravitational waves from known pulsars with science run 5 LIGO data. *The Astrophysical Journal*, 713(1):671.
- [14] Akgün, T., Link, B., and Wasserman, I. (2006). Precession of the isolated neutron star PSR B1828-11. *Monthly Notices of the Royal Astronomical Society*, 365:653–672.
- [15] Alpar, M. A., Pines, D., Anderson, P. W., and Shaham, J. (1984). Vortex creep and the internal temperature of neutron stars. i-general theory. *The Astrophysical Journal*, 276:325–334.
- [16] Anderson, P. W. and Itoh, N. (1975). Pulsar glitches and restlessness as a hard superfluidity phenomenon. *Nature*.
- [17] Andersson, N. (2003). Gravitational waves from instabilities in relativistic stars. *Classical and Quantum Gravity*, 20(7):R105.
- [18] Andersson, N., Glampedakis, K., Ho, W. C. G., and Espinoza, C. M. (2012). Pulsar glitches: The crust is not enough. *Physical Review Letters*, 109:241103.
- [19] Andersson, N. and Kokkotas, K. D. (2001). The r-mode instability in rotating neutron stars. *International Journal of Modern Physics D*, 10(04):381–441.
- [20] Archibald, A. M., Stairs, I. H., Ransom, S. M., Kaspi, V. M., Kondratiev, V. I., Lorimer, D. R., McLaughlin, M. A., Boyles, J., Hessels, J. W., Lynch, R., et al. (2009). A radio pulsar/X-ray binary link. *Science*, 324(5933):1411–1414.
- [21] Archibald, R. F., Kaspi, V. M., Ng, C.-Y., Gourgouliatos, K. N., Tsang, D., Scholz, P., Beardmore, A. P., Gehrels, N., and Kennea, J. A. (2013). An anti-glitch in a magnetar. *Nature*, 497(7451):591–593.

- [22] Arzamasskiy, L., Philippov, A., and Tchekhovskoy, A. (2015). Evolution of non-spherical pulsars with plasma-filled magnetospheres. *Monthly Notices of the Royal Astronomical Society*, 453:3540–3553.
- [23] Arzoumanian, Z., Nice, D. J., Taylor, J. H., and Thorsett, S. E. (1994). Timing behavior of 96 radio pulsars. *The Astrophysical Journal*, 422:671–680.
- [24] Ashton, G., Jones, D. I., and Prix, R. (2015). Effect of timing noise on targeted and narrow-band coherent searches for continuous gravitational waves from pulsars. *Physical Review D*, 91(6):062009.
- [25] Ashton, G., Jones, D. I., and Prix, R. (2016). Comparing models of the periodic variations in spin-down and beamwidth for PSR B1828-11. *Monthly Notices of the Royal Astronomical Society*, 458:881–899.
- [26] Baade, W. and Zwicky, F. (1934). Cosmic Rays from Super-novae. *Proceedings of the National Academy of Science*, 20:259–263.
- [27] Bailes, M., Lyne, A. G., and Shemar, S. L. (1993). Limits on pulsar planetary systems from the Jodrell Bank timing database. In Phillips, J. A., Thorsett, S. E., and Kulkarni, S. R., editors, *Planets Around Pulsars*, volume 36 of *Astronomical Society of the Pacific Conference Series*, pages 19–30.
- [28] Baym, G. and Pines, D. (1971). Neutron starquakes and pulsar speedup. *Annals of Physics*, 66(2):816–835.
- [29] Biryukov, A., Beskin, G., and Karpov, S. (2012). Monotonic and cyclic components of radio pulsar spin-down. *Monthly Notices of the Royal Astronomical Society*, 420(1):103–117.
- [30] Bisnovatyi-Kogan, G. S., Mersov, G. A., and Sheffer, E. K. (1990). Model of the 35-day cycle in the X-Ray Binary Hercules X-1. *Soviet Astronomy*, 34:44.
- [31] Bisnovatyi-Kogan, G. and Kahabka, P. (1993). Period variations and phase residuals in freely precessing stars. *Astronomy & Astrophysics*, 267:L43–L46.
- [32] Bombaci, I. (1996). The maximum mass of a neutron star. *Astronomy & Astrophysics*, 305:871.
- [33] Boynton, P., Groth, E., Hutchinson, D., Nanos Jr, G., Partridge, R., and Wilkinson, D. (1972). Optical timing of the Crab pulsar, NP 0532. *The Astrophysical Journal*, 175:217.
- [34] Boynton, P. E., Groth, III, E. J., Partridge, R. B., and Wilkinson, D. T. (1969). Apparent change in frequency of NP 0532. *IAU Circ.*, 2179.
- [35] Brady, P. R., Creighton, T., Cutler, C., and Schutz, B. F. (1998). Searching for periodic sources with LIGO. *Physical Review D*, 57:2101–2116.
- [36] Brits, L. (2010). Euler angles. <https://commons.wikimedia.org/wiki/File:Eulerangles-alternative.svg>.

- [37] Burgay, M., D'Amico, N., Possenti, A., Manchester, R., Lyne, A., Joshi, B., McLaughlin, M., Kramer, M., Sarkissian, J., Camilo, F., et al. (2003). An increased estimate of the merger rate of double neutron stars from observations of a highly relativistic system. *Nature*, 426(6966):531–533.
- [38] Chadwick, J. (1932). The existence of a neutron. *Proceedings of the Royal Society of London. Series A*, 136(830):692–708.
- [39] Chamel, N. and Haensel, P. (2008). Physics of neutron star crusts. *Living Rev. Relativity*, 11(10).
- [40] Cheng, K. (1987). Outer magnetospheric fluctuations and pulsar timing noise. *The Astrophysical Journal*.
- [41] Chiuderi, C. and Occhionero, F. (1970). Shape of the Crab Pulsar and its Period Fluctuations. *Nature*, 226:337–338.
- [42] Coles, W., Hobbs, G., Champion, D. J., Manchester, R. N., and Verbiest, J. P. W. (2011). Pulsar timing analysis in the presence of correlated noise. *Monthly Notices of the Royal Astronomical Society*, 418:561–570.
- [43] Cordes, J. M. (1980). Pulsar timing. II-Analysis of random walk timing noise- Application to the Crab pulsar. *The Astrophysical Journal*, 237:216–226.
- [44] Cordes, J. M. (1993). The detectability of planetary companions to radio pulsars. In *Planets around pulsars*, pages 43–60.
- [45] Cordes, J. M. (2013). Pulsar State Switching from Markov Transitions and Stochastic Resonance. *ApJ*, 775:47.
- [46] Cordes, J. M. and Downs, G. S. (1985). JPL pulsar timing observations. III-Pulsar rotation fluctuations. *The Astrophysical Journal Supplement Series*, 59:343–382.
- [47] Cordes, J. M. and Greenstein, G. (1981). Pulsar timing. IV-Physical models for timing noise processes. *The Astrophysical Journal*, 245:1060–1079.
- [48] Cordes, J. M. and Helfand, D. J. (1980). Pulsar timing. iii-timing noise of 50 pulsars. *The Astrophysical Journal*, 239:640–650.
- [49] Cordes, J. M. and Shannon, R. M. (2008). Rocking the lighthouse: circumpulsar asteroids and radio intermittency. *The Astrophysical Journal*, 682(2):1152.
- [50] Craft, Jr., H. D. (1970). *Radio Observations of the Pulse Profiles and Dispersion Measures of Twelve Pulsars*. PhD thesis, CORNELL UNIVERSITY.
- [51] d'Agostino, R. B. (1971). An omnibus test of normality for moderate and large size samples. *Biometrika*, 58(2):341–348.
- [52] Davis, L. and Goldstein, M. (1970). Magnetic-dipole alignment in pulsars. *The Astrophysical Journal*, 159.
- [53] Dergachev, V. (2010). On blind searches for noise dominated signals: a loosely coherent approach. *Classical and Quantum Gravity*, 27(20):205017.

- [54] Deutsch, A. (1955). The electromagnetic field of an idealized star in rigid rotation in vacuo. *Annales d'Astrophysique*.
- [55] Dodson, R. G., McCulloch, P. M., and Lewis, D. R. (2001). High time resolution observations of the January 2000 glitch in the Vela pulsar. *The Astrophysical Journal Letters*, 564(2):L85.
- [56] Duncan, R. C. and Thompson, C. (1996). Magnetars. In *High velocity neutron stars and gamma-ray bursts*, volume 366, pages 111–117. AIP Publishing.
- [57] Edwards, R. T., Hobbs, G. B., and Manchester, R. N. (2006). TEMPO2, a new pulsar timing package - II. The timing model and precision estimates. *Monthly Notices of the Royal Astronomical Society*, 372:1549–1574.
- [58] Einstein, A. (1916). Approximative integration of the field equations of gravitation. *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)*, 688:1916.
- [59] Espinoza, C. M., Lyne, A. G., Stappers, B. W., and Kramer, M. (2011). A study of 315 glitches in the rotation of 102 pulsars. *Monthly Notices of the Royal Astronomical Society*, 414:1679–1704.
- [60] Faucher-Giguere, C.-A. and Kaspi, V. M. (2006). Birth and evolution of isolated radio pulsars. *The Astrophysical Journal*, 643(1):332.
- [61] Foreman-Mackey, D. (2016). corner.py: Scatterplot matrices in Python. *The Journal of Open Source Software*, 24.
- [62] Foreman-Mackey, D., Hogg, D. W., Lang, D., and Goodman, J. (2013). emcee: The MCMC Hammer. *Publications of the Astronomical Society of the Pacific*, 125:306–312.
- [63] Gavriil, F. P., Kaspi, V. M., and Woods, P. M. (2002). Magnetar-like X-ray bursts from an anomalous X-ray pulsar. *Nature*, 419:142–144.
- [64] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2013). *Bayesian Data Analysis*. CRC press.
- [65] Gelman, A. et al. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian analysis*, 1(3):515–534.
- [66] Glampedakis, K. and Jones, D. I. (2010). Implications of magnetar non-precession. *Monthly Notices of the Royal Astronomical Society: Letters*, 405(1):L6–L10.
- [67] Goggans, P. M. and Chi, Y. (2004). Using thermodynamic integration to calculate the posterior probability in bayesian model selection problems. *AIP Conference Proceedings*, 707(1):59–66.
- [68] Gold, T. (1968). Rotating Neutron Stars as the Origin of the Pulsating Radio Sources. *Nature*, 218:731–732.
- [69] Goldreich, P. (1970). Neutron star crusts and alignment of magnetic axes in pulsars. *The Astrophysical Journal*.

- [70] Goldreich, P. and Julian, W. H. (1969). Pulsar electrodynamics. *The Astrophysical Journal*.
- [71] Goodman, J. and Weare, J. (2012). Ensemble samplers with affine invariance. *Comm. App. Math. Comp. Sci.*, 5.
- [72] Gough, B. (2009). *GNU scientific library reference manual*. Network Theory Ltd.
- [73] Groth, E. J. (1975). Timing of the Crab pulsar III. The slowing down and the nature of the random process. *The Astrophysical Journal Supplement Series*, 29:453–465.
- [74] Gunn, J. E. and Ostriker, J. P. (1969). Acceleration of high-energy cosmic rays by pulsars. *Physical Review Letters*, 22(14):728.
- [75] Hawking, S. W. and Israel, W. (1989). *Three Hundred Years of Gravitation*. Cambridge University Press.
- [76] Helfand, D. J., Taylor, J. H., Backus, P. R., and Cordes, J. M. (1980). Pulsar timing. i-observations from 1970 to 1978. *The Astrophysical Journal*, 237:206–215.
- [77] Hellings, R. W. and Downs, G. S. (1983). Upper limits on the isotropic gravitational radiation background from pulsar timing analysis. *The Astrophysical Journal*, 265:L39–L42.
- [78] Hewish, A., Bell, S. J., Pilkington, J. D. H., Scott, P. F., and Collins, R. A. (1968). Observation of a Rapidly Pulsating Radio Source. *Nature*, 217:709–713.
- [79] Hobbs, G., Archibald, A., Arzoumanian, Z., Backer, D., Bailes, M., Bhat, N. D. R., Burgay, M., Burke-Spolaor, S., Champion, D., Cognard, I., et al. (2010a). The international pulsar timing array project: using pulsars as a gravitational wave detector. *Classical and Quantum Gravity*, 27(8):084013.
- [80] Hobbs, G., Edwards, R., and Manchester, R. (2006). TEMPO2: a New Pulsar Timing Package. *Chinese Journal of Astronomy and Astrophysics Supplement*, 6(2):189–192.
- [81] Hobbs, G., Lyne, A. G., and Kramer, M. (2010b). An analysis of the timing irregularities for 366 pulsars. *Monthly Notices of the Royal Astronomical Society*, 402(2):1027–1048.
- [82] Hu, Y.-M., Pitkin, M., Heng, I. S., and Hendry, M. A. (2014). Glitch or anti-glitch: a Bayesian view. *The Astrophysical Journal Letters*, 784(2):L41.
- [83] Hulse, R. A. and Taylor, J. H. (1975). Discovery of a pulsar in a binary system. *The Astrophysical Journal*, 195:L51–L53.
- [84] Janssen, G. H. and Stappers, B. W. (2006). 30 glitches in slow pulsars. *Astronomy & Astrophysics*, 457:611–618.
- [85] Jaranowski, P., Królak, A., and Schutz, B. F. (1998). Data analysis of gravitational-wave signals from spinning neutron stars: The signal and its detection. *Physical Review D*, 58(6):063001.

- [86] Jaynes, E. T. (2003). *Probability theory – the logic of science*. Cambridge university press.
- [87] Jeffreys, H. (1998). *The theory of probability*. OUP Oxford.
- [88] Jones, D. I. (2004). Is timing noise important in the gravitational wave detection of neutron stars? *Physical Review D*, 70(4):1–9.
- [89] Jones, D. I. (2012). Pulsar state switching, timing noise and free precession. *Monthly Notices of the Royal Astronomical Society*, 420(3):2325–2338.
- [90] Jones, D. I. and Andersson, N. (2001). Freely precessing neutron stars: model and observations. *Monthly Notices of the Royal Astronomical Society*, 324(4):811–824.
- [91] Jones, D. I. and Andersson, N. (2002). Gravitational waves from freely precessing neutron stars. *Monthly Notices of the Royal*
- [92] Jones, E., Oliphant, T., Peterson, P., et al. (2001). SciPy: Open source scientific tools for Python. [Online; accessed 2015-11-26].
- [93] Jones, P. B. (1988). Excitation of small-amplitude free precession in the Crab pulsar. *Monthly Notices of the Royal Astronomical Society*, 235(2):545–550.
- [94] Jung, P. (1991). Amplification of small signals via stochastic resonance. *Physical Review A*, 44(12).
- [95] Kaspi, V. M. (1996). Pulsar/Supernova Remnant Associations. In Johnston, S., Walker, M. A., and Bailes, M., editors, *IAU Colloq. 160: Pulsars: Problems and Progress*, volume 105 of *Astronomical Society of the Pacific Conference Series*, page 375.
- [96] Kouveliotou, C., Duncan, R. C., and Thompson, C. (2003). Magnetars. *Scientific American*, 288(2):34–41.
- [97] Kramer, M., Lyne, A. G., O’Brien, J. T., Jordan, C. A., and Lorimer, D. R. (2006a). A Periodically Active Pulsar Giving Insight into Magnetospheric Physics. *Science*, 312:549–551.
- [98] Kramer, M., Stairs, I. H., Manchester, R. N., McLaughlin, M. A., Lyne, A. G., Ferdman, R. D., Burgay, M., Lorimer, D. R., Possenti, A., D’Amico, N., Sarkissian, J. M., Hobbs, G. B., Reynolds, J. E., Freire, P. C. C., and Camilo, F. (2006b). Tests of General Relativity from Timing the Double Pulsar. *Science*, 314:97–102.
- [99] Krishnan, B., Sintes, A. M., Papa, M. A., Schutz, B. F., Frasca, S., and Palomba, C. (2004). Hough transform search for continuous gravitational waves. *Physical Review D*, 70(8):082001.
- [100] Landau, L. D. and Lifshitz, E. M. (1969). *Mechanics*, volume 1. Pergamon press, second edition.
- [101] Landau, L. D. and Lifshitz, E. M. (1971). *The classical theory of fields*, volume 2. Pergamon press, third revised english edition.

- [102] Leaci, P. and Prix, R. (2015). Directed searches for continuous gravitational waves from binary systems: Parameter-space metrics and optimal Scorpius X-1 sensitivity. *Physical Review D*, 91(10):102003.
- [103] Lee, K. J., Wex, N., Kramer, M., Stappers, B. W., Bassa, C. G., Janssen, G. H., Karuppusamy, R., and Smits, R. (2011). Gravitational wave astronomy of single sources with a pulsar timing array. *Monthly Notices of the Royal Astronomical Society*, 414(4):3251–3264.
- [104] Lentati, L., Alexander, P., Hobson, M. P., Feroz, F., van Haasteren, R., Lee, K. J., and Shannon, R. M. (2014). TEMPONEST: a Bayesian approach to pulsar timing analysis. *Monthly Notices of the Royal Astronomical Society*, 437:3004–3023.
- [105] Lewin, W. H., van Paradijs, J., and van den Heuvel, E. P. (1997). *X-ray Binaries*, volume 26. Cambridge University Press.
- [106] LIGO Scientific Collaboration (2014). LALSuite: FreeSoftware (GPL) Tools for Data-Analysis.
- [107] Link, B. (2003). Constraining Hadronic Superfluidity with Neutron Star Precession. *Physical Review Letters*, 91(10):101101.
- [108] Link, B. and Epstein, R. I. (2001). Precession interpretation of the isolated pulsar PSR B1828-11. *The Astrophysical Journal*, 556(1):392.
- [109] Lyne, A. (2012). Timing noise and the long-term stability of pulsar profiles. *Proceedings of the International Astronomical Union*, 8(S291):183–188.
- [110] Lyne, A. and Graham-Smith, F. (2012). *Pulsar Astronomy*. Cambridge University Press, fourth edition.
- [111] Lyne, A., Hobbs, G., Kramer, M., Stairs, I., and Stappers, B. (2010). Switched magnetospheric regulation of pulsar spin-down. *Science*, 329(5990):408–412.
- [112] Lyne, A. G. and Manchester, R. N. (1988). The shape of pulsar radio beams. *Monthly Notices of the Royal Astronomical Society*, 234:477–508.
- [113] Lyne, A. G., Pritchard, R. S., and Graham-Smith, F. (1993). Twenty-Three Years of Crab Pulsar Rotational History. *Monthly Notices of the Royal Astronomical Society*, 265:1003.
- [114] Lyne, A. G., Stairs, I. H., and Shemar, S. L. (2000). Periodicities in Rotation and Pulse Shape in PSR B1828-11. In Kramer, M., Wex, N., and Wielebinski, R., editors, *IAU Colloq. 177: Pulsar Astronomy - 2000 and Beyond*, volume 202 of *Astronomical Society of the Pacific Conference Series*, page 93.
- [115] Maciesiak, K., Gil, J., and Ribeiro, V. A. R. M. (2011). On the pulse-width statistics in radio pulsars - I. Importance of the interpulse emission. *Monthly Notices of the Royal Astronomical Society*, 414:1314–1328.
- [116] MacKay, D. J. C. (2003). *Information theory, inference and learning algorithms*. Cambridge university press.

- [117] Manchester, R. N., Hobbs, G. B., Teoh, A., and Hobbs, M. (2005). The Australia Telescope National Facility Pulsar Catalogue. *The Astronomical Journal*, 129:1993–2006.
- [118] Manchester, R. N. and Lyne, A. G. (1977). Pulsar interpulses - Two poles or one. *Monthly Notices of the Royal Astronomical Society*, 181:761–767.
- [119] Matsakis, D. N., Taylor, J. H., and Eubanks, T. M. (1997). A statistic for describing pulsar and clock stabilities. *Astronomy & Astrophysics*, 326:924–928.
- [120] McKenna, J. and Lyne, A. G. (1990). PSR1737 - 30 and period discontinuities in young pulsars. *Nature*, 343:349.
- [121] Melatos, A. (1999). Bumpy spin-down of anomalous X-ray pulsars: The link with magnetars. *The Astrophysical Journal Letters*, 519(1):L77.
- [122] Melatos, A. (2000). Radiative precession of an isolated neutron star. *Monthly Notices of the Royal Astronomical Society*, 313(2):217–228.
- [123] Melatos, A., Peralta, C., and Wyithe, J. S. B. (2008). Avalanche dynamics of radio pulsar glitches. *The Astrophysical Journal*, 672(2):1103.
- [124] Messenger, C., Bulten, H. J., Crowder, S. G., et al. (2014). Gravitational waves from Sco X-1: A comparison of search methods. *in preparation.* (LIGO DCC-P1400217).
- [125] Michel, F. C. and Goldwire, Jr., H. C. (1970). Alignment of Oblique Rotators. In *Bulletin of the American Astronomical Society*, volume 2 of *Bulletin of the American Astronomical Society*, page 209.
- [126] Mitton, S. (1977). *The Cambridge encyclopaedia of astronomy*. Jonathan Cape Ltd.
- [127] Narayan, R. and Vivekanand, M. (1983). Evidence for evolving elongated pulsar beams. *Astronomy & Astrophysics*, 122:45–53.
- [128] Nelson, R. W., Finn, L. S., and Wasserman, I. (1990). Trompe l’oeil’binary’pulsars. *The Astrophysical Journal*, 348:226–231.
- [129] O’Hagan, A. and Leonard, T. (1976). Bayes estimation subject to uncertainty about parameter constraints. *Biometrika*, 63(1):201–203.
- [130] Oppenheimer, J. R. and Volkoff, G. M. (1939). On massive neutron cores. *Physical Review*, 55(4):374.
- [131] Pacini, F. (1967). Energy Emission from a Neutron Star. *Nature*, 216:567–568.
- [132] Pacini, F. (1968). Rotating neutron stars, pulsars and supernova remnants. *Nature*, 219:145–146.
- [133] Pavlov, G. G. and Zavlin, V. E. (2003). Thermal radiation from cooling neutron stars. In *Texas in Tuscany. XXI Texas Symposium on Relativistic Astrophysics*, pages 319–328.

- [134] Perera, B. B. P., Stappers, B. W., Weltevrede, P., Lyne, A. G., and Bassa, C. G. (2015). Understanding the spin-down rate changes of PSR B0919+06. *Monthly Notices of the Royal Astronomical Society*, 446:1380–1388.
- [135] Perera, B. B. P., Stappers, B. W., Weltevrede, P., Lyne, A. G., and Rankin, J. M. (2016). Correlated spin-down rates and radio emission in PSR B1859+07. *Monthly Notices of the Royal Astronomical Society*, 455:1071–1078.
- [136] Pines, D. and Shaham, J. (1972). The elastic energy and character of quakes in solid stars and planets. *Physics of the Earth and Planetary Interiors*, 6(1):103–115.
- [137] Pitkin, M. and Woan, G. (2004). Searching for gravitational waves from the Crab pulsar—the problem of timing noise. *Classical and Quantum Gravity*, 21:843.
- [138] Popov, S. B., Pons, J. A., Miralles, J. A., Boldin, P. A., and Posselt, B. (2010). Population synthesis studies of isolated neutron stars with magnetic field decay. *Monthly Notices of the Royal Astronomical Society*, 401(4):2675–2686.
- [139] Prix, R. (2007). Search for continuous gravitational waves: metric of the multide-tector F-statistic. *Physical Review D*, 75(2):023004.
- [140] Prix, R. (2009). Gravitational Waves from Spinning Neutron Stars. In Becker, W., editor, *Astrophysics and Space Science Library*, volume 357 of *Astrophysics and Space Science Library*, page 651. <https://dcc.ligo.org/LIGO-P060039/public>.
- [141] Prix, R. and Itoh, Y. (2005). Global parameter-space correlations of coherent searches for continuous gravitational waves. *Classical and Quantum Gravity*, 22:1003.
- [142] Prix, R. and Shaltev, M. (2012). Search for continuous gravitational waves: Optimal StackSlide method at fixed computing cost. *Physical Review D*, 85(8):084010.
- [143] Radhakrishnan, V. and Cooke, D. J. (1969). Magnetic Poles and the Polarization Structure of Pulsar Radiation. *Astrophysical Letters*, 3:225.
- [144] Radhakrishnan, V. and Manchester, R. N. (1969). Detection of a Change of State in the Pulsar PSR 0833-45. *Nature*, 222:228–229.
- [145] Reichley, P. E. and Downs, G. S. (1969). Observed Decrease in the Periods of Pulsar PSR 0833-45. *Nature*, 222:229–230.
- [146] Richards, D. W., Pettengill, G. H., Counselman, C. C., and Rankin, J. (1969a). Quasi-sinusoidal components in arrival time of pulsar NP 0532. *IAU Circ.*, 2178.
- [147] Richards, D. W., Pettengill, G. H., Roberts, J. A., Counselman, C. C., and Rankin, J. (1969b). Np 0532. *IAU Circ.*, 2181.
- [148] Ruderman, M. (1969). Neutron Starquakes and Pulsar Periods. *Nature*, 223:597–598.
- [149] Ruderman, M. (1970). Pulsar wobble and neutron starquakes. *Nature*, 225:838–839.

- [150] Sedrakian, A., Wasserman, I., and Cordes, J. M. (1999). Precession of isolated neutron stars. I. Effects of imperfect pinning. *The Astrophysical Journal*, 524(1):341.
- [151] Seymour, A. D. and Lorimer, D. R. (2013). Evidence for chaotic behaviour in pulsar spin-down rates. *Monthly Notices of the Royal Astronomical Society*, 428:983–998.
- [152] Shaham, J. (1977). Free precession of neutron stars – role of possible vortex pinning. *The Astrophysical Journal*, 214:251–260.
- [153] Shaltev, M. and Prix, R. (2013). Fully coherent follow-up of continuous gravitational-wave candidates. *Physical Review D*, 87(8):084057.
- [154] Shapiro, S. L. and Teukolsky, S. A. (1983). *Black Holes, White Dwarfs, and Neutron Stars*. John Wiley & Sons, Inc.
- [155] Sivia, D. S. and Skilling, J. (1996). *Data analysis: a Bayesian tutorial*. Oxford university press.
- [156] Stairs, I., Lyne, A., and Shemar, S. (2000). Evidence for free precession in a pulsar. *Nature*, 406(6795):484–6.
- [157] Stairs, I. H., Athanasiadis, D., Kramer, M., and Lyne, A. G. (2003). High-Resolution Observations of PSR B1828-11. In Bailes, M., Nice, D. J., and Thorsett, S. E., editors, *Radio Pulsars*, volume 302 of *Astronomical Society of the Pacific Conference Series*, page 249.
- [158] Swendsen, R. H. and Wang, J. (1986). Replica Monte Carlo Simulation of Spin-Glasses. *Physical Review Letters*, 57:2607–2609.
- [159] Taylor, J. H. and Weisberg, J. M. (1982). A new test of general relativity-Gravitational radiation and the binary pulsar PSR 1913+ 16. *The Astrophysical Journal*, 253:908–920.
- [160] van den Bergh, S. (1991). Galactic and extragalactic supernova rates. In *Supernovae*, pages 711–719. Springer.
- [161] Wald, A. (1944). On cumulative sums of random variables. *The Annals of Mathematical Statistics*, 15(3):283–296.
- [162] Wang, J., Wang, N., Tong, H., and Yuan, J. (2012). Recent glitches detected in the Crab pulsar. *Astrophysics and Space Science*, 340(2):307–315.
- [163] Wang, N., Manchester, R. N., Pace, R. T., Bailes, M., Kaspi, V. M., Stappers, B. W., and Lyne, A. G. (2000). Glitches in southern pulsars. *Monthly Notices of the Royal Astronomical Society*, 317:843–860.
- [164] Wang, N., Wu, X.-J., Manchester, R., Zhang, J., Lyne, A., and Yusup, A. (2001). A large glitch in the Crab pulsar. *Chinese Journal of Astronomy and Astrophysics*, 1(3):195.

- [165] Watts, A. L., Krishnan, B., Bildsten, L., and Schutz, B. F. (2008). Detecting gravitational wave emission from the known accreting neutron stars. *Monthly Notices of the Royal Astronomical Society*, 389(2):839–868.
- [166] Weiss, N. A., Holmes, P. T., and Hardy, M. (2006). *A course in probability*. Pearson Addison Wesley Boston, Massachusetts, USA.
- [167] Weisstein, E. W. (2014). Least squares fitting-polynomial. From MathWorld-A Wolfram Web Resource. <http://mathworld.wolfram.com/LeastSquaresFittingPolynomial.html>, last visited on 8/9/2014.
- [168] Weisstein, E. W. (2016). Normal difference distribution. From MathWorld-A Wolfram Web Resource. <http://mathworld.wolfram.com/NormalDifferenceDistribution.html>, last visited on 15/04/2016.
- [169] Wijnands, R. and Van der Klis, M. (1998). A millisecond pulsar in an X-ray binary system. *nature*, 394(6691):344–346.
- [170] Wolszczan, A. and Frail, D. A. (1992). A planetary system around the millisecond pulsar psr 1257+12. *Nature*, 355(6356):145–147.
- [171] Yakovlev, D. G., Haensel, P., Baym, G., and Pethick, C. (2013). Lev Landau and the concept of neutron stars. *Physics Uspekhi*, 56:289–295.