The Pennsylvania State University

The Graduate School

Department of Physics

BAYESIAN ANALYSIS ON GRAVITATIONAL WAVES AND EXOPLANETS

A Dissertation in

Physics

by

Xihao Deng

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2015

The dissertation of Xihao Deng was read and approved[1] by the following:

Lee Samuel Finn
Professor of Physics
Dissertation Advisor
Chair of Committee

Chad Hanna
Assistant Professor of Physics

Stephane Coutu
Professor of Physics

Aleksander Wolszczan
Professor of Astronomy and Astrophysics

Nitin Samarth
Professor of Physics
Head of the Department of Physics

[1]Signatures on file in the Graduate School.

Abstract

Attempts to detect gravitational waves using a pulsar timing array (PTA), i.e., a collection of pulsars in our Galaxy, have become more organized over the last several years. PTAs act to detect gravitational waves generated from very distant sources by observing the small and correlated effect the waves have on pulse arrival times at the Earth. In this thesis, I present advanced Bayesian analysis methods that can be used to search for gravitational waves in pulsar timing data. These methods were also applied to analyze a set of radial velocity (RV) data collected by the Hobby-Eberly Telescope on observing a K0 giant star. They confirmed the presence of two Jupiter mass planets around a K0 giant star and also characterized the stellar p-mode oscillation.

The first part of the thesis investigates the effect of wavefront curvature on a pulsar's response to a gravitational wave. In it we show that we can assume the gravitational wave phasefront is planar across the array only if the source luminosity distance $\gg 2\pi L^2/\lambda$, where L is the pulsar distance to the Earth ($\sim$ kpc) and λ is the radiation wavelength ($\sim$ pc) in the PTA waveband. Correspondingly, for a point gravitational wave source closer than ~ 100 Mpc, we should take into account the effect of wavefront curvature across the pulsar-Earth line of sight, which depends on the luminosity distance to the source, when evaluating the pulsar timing response. As a consequence, if a PTA can detect a gravitational wave from a source closer than $\sim$ 100 Mpc, the effects of wavefront curvature on the response allows us to determine the source luminosity distance. This technique is very similar to the use of pulsar timing parallax to measure the distances to nearby pulsars, because they both try to measure the phasefront curvature of a wave passing through a long baseline (pulsar-Earth distance and Sun-Earth distance) to determine the wave source distance.

The second and third parts of the thesis propose a new analysis method based on Bayesian nonparametric regression to search for gravitational wave bursts and a gravitational wave background in PTA data. Unlike the conventional Bayesian analysis that introduces a signal model with a fixed number of parameters, Bayesian nonparametric regression sets constraints on the *function space* that may be reasonably thought to characterize the range of gravitational wave signals. For example, focus attention on the detection of a gravitational wave burst, by which we mean a signal that begins and ends over the course of an observational epoch. The burst may result from a source that we know how to model - e.g., a near-unity mass ratio black hole binary system - or it may be the result of a process, which we have not imagined and, so, have no model for. Similarly, a gravitational wave background resulting from a superposition of a number of weak sources may be difficult to characterize if the number of weak sources is sufficiently large that none can be individually resolved, but not so large that their superposition leads to a reasonably Gaussian distribution. Correspondingly, the Bayesian nonparametric regression method may be very useful to help search for gravitational wave bursts and a gravitational wave background in the pulsar timing data. By testing this new method on simulated data sets, it is found that we can use it to detect gravitational wave bursts and a gravitational wave background, and we can also characterize their important physical parameters such as the burst durations and the amplitude of the background even if their signal-to-noise ratios are low.

The fourth part develops Bayesian analysis methods that can be used to detect gravitational waves generated from circular-orbit supermassive black hole binaries with a pulsar timing array. PTA response to such gravitational waves can be modeled as the difference between two sinusoidal terms — the one with a coherent phase among different pulsars called “Earth term” and the other one with incoherent phases among different pulsars called “pulsar term”. For gravitational waves from slowly evolving binaries, the two terms in the PTA response model have the same frequency.

Previous methods aimed at detecting gravitational waves from circular-orbit binaries ignored pulsar terms in data analysis since those terms were considered to be negligible when averaging over all the pulsars. However, it is found that we can incorporate the contributions of pulsar terms into data analysis in the case of slowly evolving binaries by treating the incoherent phases in pulsar terms as unknown parameters to be marginalized. By testing the new method on simulated data sets, the improvement, compared to previous analyses, is equivalent to halving the strength of timing noise associated with each pulsar.

The final part of this thesis applies Bayesian analysis to search for the evidence of a planetary system around the K0 giant star HD 102103 detected by the Penn State-Torun planet group at the Hobby Eberly Telescope. It analyzes 116 observations of the star's radial velocity. However, the stellar p-mode oscillation also contributes to the radial velocity data, challenging the search for the planets around the star. The Bayesian method models the stellar oscillation effect and the potential exoplanet signal together, simultaneously inferring their parameters from the data. Consequently, the method removes the ambiguities of the presence of two Jupiter mass planets around the K0 giant and as a bonus, it also characterizes the strength and the frequency of the stellar oscillation.

Table of Contents

List of Tables

List of Figures

Acknowledgments

I would like to thank my advisor Prof. Lee Samuel Finn for supporting me to work on all of these projects, and giving me a lot of valuable suggestions on research and academic writing. I would like to thank Prof. Alex Wolszczan for offering me the radial velocity data of HD 1020103 and helping me to analyze the data. I would like to thank Prof. Michael Eracleous for giving me valuable guidance on the formation and evolution of supermassive black hole binaries and Prof. Peter Meszaros for his fruitful instructions on gamma ray bursts. I would like to thank the National Science Foundation and the Penn State Physics Department for financial support and other research awards including the Duncan Fellowships, Downsbrough Fellowship, Mebus Fellowship, W. Donald Miller Fellowship and Eklund Lecture Award. Finally, I would like to thank my parents for priceless love and support.

Chapter 1

Introduction

It has been 30 years since Sazhin and Detweiler showed how gravitational waves could be directly detected by correlating the timing data of a collection of pulsars (Sazhin, 1978; Detweiler, 1979). Over the ensuing years, attempts to detect gravitational waves with a pulsar timing array, i.e., a collection of pulsars, have developed rapidly as pulsar timing techniques have drastically progressed. Such an "astronomical" detector is sensitive to gravitational waves with frequency in the nHz range, the sources of which include supermassive black hole binaries or encounters resulting from galaxy mergers (Sesana et al., 2008, 2009). Therefore, successful detections of gravitational waves with a pulsar timing array can help study the processes of galaxy mergers and the evolutions of double supermassive black hole systems. In this thesis, I mainly describe several data analysis methods I have developed and tested in simulated pulsar timing data sets. I will also describe an application of these methods to the radial velocity data collected by the Hobby Eberly Telescope to confirm the evidence of two Jupiter mass exoplanets around a red giant star.

1.1 Overview of pulsar timing

Pulsars emit continuous beams of radiation. The beams rotate with the pulsar. We observe a pulse of radiation when the beam intersects the pulsar-Earth line-of-sight. Pulsar timing array techniques are measuring the arrival times of the radio pulses (TOAs) with large radio telescopes such as Arecibo, Green Bank, Parkes, etc (Hobbs et al., 2010), because timing the radio pulses obtains much better timing precision than timing pulses of other electromagnetic wavelengths. These TOAs need adjustments so that they represent the arrival times in an inertial frame. This is

accomplished by transforming each measured TOA to an arrival time at the solar system barycenter (SSB) and adding additional terms that can model the other physical effects influencing TOAs. Such a pulsar timing model can be written as (Hobbs et al., 2006):

$$t = \mathrm{TOA} + \Delta_{\mathrm{C}} + \Delta_{\mathrm{A}} + \Delta_{\mathrm{R}} + \Delta_{\mathrm{E}} + \Delta_{\mathrm{S}} - \mathrm{DM}/f^2 + \Delta_{\mathrm{VP}} + \Delta_{\mathrm{B}} \tag{1.1}$$

where Δ_{C} contains various clock corrections (see 1.1.1); Δ_{A} denotes the atmospheric propagation delays (see 1.1.2); Δ_{R} denotes the solar system Roemer delay (see 1.1.3); Δ_{E} denotes the solar system Einstein delay (see 1.1.4); Δ_{S} denotes the solar system Shapiro delay (see 1.1.4); DM/f^2 models the dispersive component of the light travel time (see 1.1.5); Δ_{VP} describes the excess vacuum propagation delay due to secular motion (see 1.1.6); and Δ_{B} denotes the effects induced by the orbital motion of the pulsar if the pulsar is in a binary system (see 1.1.7). Table 1.1 lists the typical magnitude of the primary timing corrections (Hobbs et al., 2006; Edwards et al., 2006).

Table 1.1 Typical magnitude of primary timing corrections

Timing corrections	Typical magnitude
Observatory clock correction (Δ_{C})	~ 1 μs
Atmospheric propagation delays (Δ_{A})	~ 10 ns
Roemer delay (Δ_{R})	~ 500 s
Einstein delay (Δ_{E})	~ 1 ms
Shapiro delay (Δ_{S})	~ 100 μs
Interstellar medium dispersion delay (DM/f^2)	~ 1 s
Secular motion (Δ_{VP})	~ 1 ms
Orbital motion (Δ_{B})	~ 100 ns

1.1.1 Observatory clock correction

The TOAs are recorded by the local observatory clocks, which are derived from a precision frequency standard with good short term stability, such as a hydrogen maser. On longer time scales (months to years) these clocks deviate significantly from uniformity and are therefore unsuitable for long term precision pulsar timing. However, it is generally possible to remove these errors down to the precision provided by the best available terrestrial time scale through the application of corrections derived from monitoring the offsets between pairs of clocks. The ultimate aim of the clock correction process is to transform the TOA measurements into the Geocentric Celestial Reference System, and different observatories use different transformation methods. The typical magnitude of the clock correction on TOAs is $\sim 1\mu$s (Hobbs et al., 2006; Edwards et al., 2006).

1.1.2 Atmospheric propagation delay

The group velocity of radio waves in the atmosphere differs from the vacuum speed of light. Refractivity is induced both by the ionized fraction of the atmosphere and the neutral fraction (Hobbs et al., 2006). Pulsar timing uses atmospheric pressure data to calibrate the atmospheric propagation delay which is of the order of 10 ns (Hobbs et al., 2006).

1.1.3 Roemer delay

The Roemer delay is the non-relativistic vacuum light travel time between the pulse arriving at the observatory and the equivalent arrival time at the solar system barycenter. This is calculated by determining the time delay between a pulse arriving at the observatory and the Earth center and, with the aid of the solar system ephemeris, from the Earth center to the SSB (Edwards et al., 2006).

Denoting the vector distance from the observatory to the SSB as $\vec{r}_{\rm ob}$, it can be expressed as

$$\vec{r}_{\rm ob} = \vec{r}_{\rm oe} + \vec{r}_{\rm es} + \vec{r}_{\rm sb} \tag{1.2}$$

where $\vec{r}_{\rm oe}$ denotes the vector from the observatory to the Earth center; $\vec{r}_{\rm es}$ denotes the vector from the Earth center to the center of the Sun; and $\vec{r}_{\rm sb}$ denotes the vector from the center of the Sun to the SSB. The Roemer delay would be

$$\Delta_{\rm R} = -\frac{\vec{r}_{\rm ob} \cdot \hat{s}}{c} \tag{1.3}$$

where c is the speed of light and $\hat{s}$ is the unit position vector of the pulsar that can be measured either by telescope pointing or interferometry. The typical magnitude of the Roemer delay is ~ 500s.

1.1.4 General relativistic corrections

General relativity primarily gives two correction terms in Eq. (1.2): the Einstein delay $\Delta_{\rm E}$ that quantifies the change in arrival times due to variations in clocks at the observatory induced by the gravitational potential of the Earth and Earth's motion in the solar potential, and the Shapiro delay $\Delta_{\rm S}$ that accounts for the time delay caused by the passage of the pulse through the large gravitational field of the solar system (Hobbs et al., 2006).

The Einstein delay is expressed as (Lyne & Graham-Smith, 2006, chap. 5)

$$\Delta_{\rm E} = \frac{\vec{r}_{\rm oe} \cdot \dot{\vec{r}}_{\rm es} + W_0 t_{\rm obs}}{c^2} + 1.6 \text{ ms} \left(\frac{7}{8} \sin E_\oplus + \frac{1}{2} e \sin 2E_\oplus + \frac{3}{8} e^2 \sin 3E_\oplus \right) \tag{1.4}$$

where $W_0 = 6.97 \times 10^{-10} c^2$ is the gravitational potential of the Earth; $t_{\rm obs}$ is the measured TOA; $E_\oplus$ and e are respectively the mean anomaly and the eccentricity of the Earth orbit.

The Shapiro delay of the solar system is the sum over the Shapiro delay of all the bodies in the solar system (Edwards et al., 2006)

$$\Delta_{\rm S} = \sum_j -\frac{2{\rm GM}_j}{c^3} \log|r_j|(1-\cos\theta_j) \tag{1.5}$$

where G is the gravitational constant, ${\rm M}_j$ is the mass of the jth solar system body, $|r_j|$ is the distance from jth body to the Earth and θ_j is the angle between the Earth-pulsar line-of-sight and the Earth-body line-of-sight.

1.1.5 Dispersion measure

When the pulse emitted from the pulsar propagates through the ionized interstellar medium, it experiences frequency dispersion. For small electron densities, the group velocity at which the pulse travels is (Lyne & Graham-Smith, 2006, chap. 5)

$$v = c\left(1 - \frac{n_e q_e^2}{2\pi m_e f^2}\right) \tag{1.6}$$

where n_e, q_e and m_e respectively denote the electron number density, electron charge in Gauss unit and electron mass; and f denotes the radio wave frequency. Hence the travel time T during which the pulse travels over distance L is

$$T = \int_0^L \frac{{\rm d}l}{v} = \frac{L}{c} + \frac{q_e^2}{2\pi m_e c f^2}\int_0^L n_e {\rm d}l \tag{1.7}$$

so the time delay caused by the frequency dispersion is

$$\Delta_{\rm DM} = \frac{q_e^2}{2\pi m_e c}\frac{\rm DM}{f^2} \tag{1.8}$$

where DM is the dispersion measure defined as $\int_0^L n_e {\rm d}l$.

1.1.6 Excess vacuum propagation delay due to secular motion

Pulsar timing measurements are affected by the secular motion of the pulsar relative to the SSB. The delay of the pulse arrival times caused by secular motion depends on the pulsar distance to the Earth (Lyne & Graham-Smith, 2006, chap. 4):

$$\Delta_{\rm VP} = \frac{v_T^2 T_{\rm obs}^2}{2Lc} \tag{1.9}$$

where v_T denotes the transverse velocity of the pulsar and $T_{\rm obs}$ denotes the observational duration. For a pulsar at distance $\sim$ 1kpc with a transverse velocity of 100 km s^{-1}, the time delay due to pulsar secular motion is of the order of 1ms.

1.1.7 Orbital motion

If the pulsar is in a binary system, its orbital motion would also influence the TOAs. The modulation of pulse arrival times by this effect is closely analogous to the cyclic variation of arrival times due to the position of the observatory relative to the SSB described above. Like the Roemer delay, we just need to transform the pulse emitting time at the pulsar to the equivalent emitting time at the center of mass of the binary system. The typical magnitude of the correction induced by pulsar orbital motion depends on the properties of the binary system and it is $\sim$ 100ns (Edwards et al., 2006).

1.1.8 Pulsar spin, spin down and timing residuals

The analysis of the timing model Eq. (1.2) helps transform the TOAs to the arrival times in the reference frame of the SSB. Now we need to consider the effects of pulsar spin and spin down. Assuming that the pulse phase at a reference time t_0 is ϕ_0, the pulse phase at an observed arrival time t_i is expressed as (Edwards et al., 2006):

$$\phi_i = \phi_0 + \nu(t_i - t_0) + \frac{1}{2}\dot{\nu}(t_i - t_0)^2 \tag{1.10}$$

where ν denotes the pulsar spin frequency and $\dot{\nu}$ is the frequency derivative which represents the spin slowing down due to the emission of energy by the pulsar. The weighted χ^2 of the timing model is (Hobbs et al., 2006)

$$\chi^2 = \sum_i \frac{1}{\sigma_i^2}\left(\frac{\phi_i - N_i}{\nu}\right)^2 \tag{1.11}$$

where N_i is the nearest integer to ϕ_i. We minimize the χ^2 to obtain the best-fit timing model. The differences between the observed TOAs and the best-fit timing model are referred to as timing residuals. The timing residuals include all other effects that are not included in Eq. (1.2) and Eq. (1.10). We expect that they may contain stochastic timing noise and gravitational wave effects that will be discussed in the next section. The rms of the current stochastic timing noise is $\sim$ 30ns $-$ 1.5μs (Demorest et al., 2012) and the amplitude of gravitational waves in the PTA waveband is $\sim 1 - 10$ns (Sesana et al., 2008, 2009).

1.2 Overview of gravitational waves

1.2.1 Gravitational wave fundamentals

The presence of gravitational waves was predicted by Einstein's general relativity and indirect evidence for its existence has been confirmed by timing the binary pulsars (Taylor et al., 1979). Here I summarize the fundamental characteristics of gravitational waves.

In general relativity, the spacetime metric $g_{\mu\nu}$ is the fundamental quantity representing the gravitation field, which obeys Einstein field equation (Misner et al., 1973):

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = \frac{16\pi \mathrm{G}}{c^4}T_{\mu\nu} \tag{1.12}$$

where $R_{\mu\nu}$ is the Ricci tensor; R is the Ricci scalar; and $T_{\mu\nu}$ is the energy momentum tensor of matter. The metric is covariant under the general coordinate transformation $x^{\mu} \rightarrow x'^{\mu}$, i.e.,

$$g'_{\mu\nu}(x') = \frac{\partial x^{\rho}}{\partial x'^{\mu}}\frac{\partial x^{\sigma}}{\partial x'^{\nu}}g_{\rho\sigma}(x) \tag{1.13}$$

where I use the Einstein summation convention. This covariant transformation is referred to as gauge transformation (Maggiore, 2007, chap. 1).

We can expand the metric around the Minkowski metric $\eta_{\mu\nu}$:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{1.14}$$

and we expand the equations of motion to linear order in $h_{\mu\nu}$ to study the linearized theory. Consider a transformation of coordinates:

$$x'^{\mu} = x^{\mu} + \xi^{\mu} \tag{1.15}$$

where the derivatives $|\partial_{\mu}\xi^{\nu}|$ are at most of the same order as $|h_{\mu\nu}|$. Using the gauge transformation law Eq. (1.13), we obtain that the transformation of $h_{\mu\nu}$, to lowest order, is

$$h'_{\mu\nu}(x') = h_{\mu\nu}(x) - (\partial_{\mu}\xi_{\nu} + \partial_{\nu}\xi_{\mu}). \tag{1.16}$$

This transformation is referred to as gauge transformation of $h_{\mu\nu}$ (Maggiore, 2007, chap. 1).

We can fix a gauge to study the independent components of $h_{\mu\nu}$. We define

$$h = \eta^{\mu\nu}h_{\mu\nu} \tag{1.17}$$

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h \tag{1.18}$$

and we use Eq. (1.16) to choose the Lorentz gauge

$$\partial^{\nu}\bar{h}_{\mu\nu} = 0. \tag{1.19}$$

Observing Eq. (1.19) does not fix the gauge completely. The Lorentz gauge in Eq. (1.19) is not influenced by a further coordinate transformation $x'^{\mu} = x^{\mu} + \xi^{\mu}$ with (Maggiore, 2007, chap. 1)

$$\Box\xi^{\mu} = 0. \tag{1.20}$$

The gauge conditions Eq. (1.19) and Eq. (1.20) fix 8 of the 10 components of $\bar{h}_{\mu\nu}$. Consequently, there are only two degrees of freedom of $\bar{h}_{\mu\nu}$.

By using Eq. (1.12), (1.14), (1.18), (1.19) and Eq. (1.20), we can obtain a simple wave equation for $\bar{h}_{\mu\nu}$:

$$\Box\bar{h}_{\mu\nu} = -\frac{16\pi\mathrm{G}}{c^4}T_{\mu\nu}. \tag{1.21}$$

This equation is similar to Maxwell's electromagnetic equation and so $\bar{h}_{\mu\nu}$ represents "gravitational waves". The two degrees of freedom of $\bar{h}_{\mu\nu}$ are referred to as the two polarization components of the gravitational waves.

To explore the detection of gravitational waves, we are interested in the wave equation Eq. (1.21) outside the gravitational wave source, i.e., where $T_{\mu\nu} = 0$,

$$\Box \bar{h}_{\mu\nu} = 0. \tag{1.22}$$

We can further simplify the form of $\bar{h}_{\mu\nu}$ in vacuum. We choose the zeroth component ξ^0 such that the trace $\bar{h} = 0$. This gauge choice does not influence the wave equations of motion since we now only consider the waves outside the source (see Eq. 1.22). Correspondingly, we have $\bar{h}_{\mu\nu} = h_{\mu\nu}$ with the traceless condition on $\bar{h}_{\mu\nu}$ (see Eq. 1.18). We then choose the three other components of ξ^μ so that $h_{0i} = 0$, where the index i ranges from 1 to 3. Consequently, the Lorentz condition Eq. (1.19) becomes

$$\begin{aligned} \partial^0 h_{00} &= 0 \\ \partial^j h_{ij} &= 0. \end{aligned} \tag{1.23}$$

We can see that h_{00} becomes a constant in time, which corresponds to the static gravitational potential of the source. We can set $h_{00} = 0$ as far as we only concern the time-dependent gravitational waves. Correspondingly, our gauge choice can be summarized as:

$$h^{0\mu} = 0, \quad h^i_i = 0, \quad \partial^j h_{ij} = 0. \tag{1.24}$$

This gauge condition is called the transverse-traceless (TT) gauge (Misner et al., 1973). The TT gauge simplifies Eq. (1.22) as (Maggiore, 2007, chap. 1)

$$\Box h_{lm}^{TT} = 0 \tag{1.25}$$

where h_{lm}^{TT} is the metric perturbation $h_{\mu\nu}$ in the TT gauge. The plane wave solution of Eq. (1.25) with gauge condition Eq. (1.24) is (Misner et al., 1973)

$$h_{lm}^{TT}(t,\vec{x}) = h_{(+)}(t-\hat{k}\cdot\vec{x})e_{lm}^{(+)}(\hat{k}) + h_{(\times)}(t-\hat{k}\cdot\vec{x})e_{lm}^{(\times)}(\hat{k}) \tag{1.26}$$

where $\hat{k}$ is the unit vector representing the gravitational wave propagation direction; $h_{(+,\times)}$ are respectively "+" and "×" polarization components of gravitational waves; $e_{lm}^{(+,\times)}$ are the polarization tensors:

$$e_{lm}^{(+)} = \hat{u}_l\hat{u}_m - \hat{v}_l\hat{v}_m \tag{1.27}$$

$$e_{lm}^{(\times)} = \hat{u}_l\hat{v}_m + \hat{v}_l\hat{u}_m \tag{1.28}$$

with $\hat{u},\hat{v}$ unit vectors orthogonal to the propagation direction $\hat{k}$ and to each other.

1.2.2 Gravitational wave detection with a pulsar timing array

Detecting gravitational waves requires being sensitive to the interactions between our detector and gravitational waves. For pulsar timing, the detector is the pulsar-Earth system and what we try to measure are the interactions between gravitational waves and the pulses generated from the pulsar. Following Finn & Lommen (2010), the jth pulsar timing response to a plane gravitational

wave can be written as

$$\tau_j(t) = \sum_{A=+,\times} F_j^{(A)}(\hat{k}) \left[\tau_{(A)}(t) - \tau_{(A)}(t - L_j(1 + \hat{k} \cdot \hat{n}_j))\right] \tag{1.29}$$

where $\tau_{(A)}$ is the integral of $h_{(A)}$

$$\frac{\mathrm{d}\,\tau_{(A)}}{\mathrm{d}\,u} = h_{(A)}(u) \tag{1.30}$$

and $F^{(A)}$ is the antenna pattern function of the pulsar-Earth system,

$$F_j^{(A)}(\hat{k}) = -\frac{\hat{n}_j^l \hat{n}_j^m e_{lm}^{(A)}(\hat{k})}{2(1 + \hat{k} \cdot \hat{n}_j)} \tag{1.31}$$

where $\hat{n}_j^l$ is the direction from Earth toward the jth pulsar, and L_j is the distance of the jth pulsar to the Earth. We can see that the pulsar timing response to a plane gravitational wave is the sum of two functionally identical terms, one time-shifted with respect to the other by an amount proportional to the Earth-pulsar distance along the wave propagation direction. The first term is referred to as the "Earth term", while the second is referred to as the "pulsar term."

Eq. (1.29) is the model for the pulse arrival time corrections induced by gravitational waves, which will be used in pulsar timing data analysis to search for evidence of gravitational waves in pulsar timing residuals. In order to reduce the false alarm probability and improve the detection sensitivity, we use a pulsar timing array, i.e., a collection of pulsars, rather than one single pulsar to detect gravitational waves, which can help search for the pattern in Eq. (1.31) that is uniquely produced by gravitational waves (Hellings & Downs, 1983; Foster & Backer, 1990).

1.2.3 Gravitational wave sources

Pulsar timing arrays are sensitive to gravitational waves of period ranging from months (observation cadence) to years (observation duration). Sources that radiate gravitational waves in this band include double black hole systems with total mass of $\sim 10^7 - 10^{10} M_\odot$ (Sesana et al., 2008, 2009).

Supermassive black holes are expected to be present at the centers of almost all massive galaxies in the universe (Haiman et al., 2009). In the late stage of galaxy mergers, supermassive black hole binaries will form or encounters between two supermassive black holes will take place (Sesana et al., 2008, 2009; Haiman et al., 2009). These are the most anticipated gravitational wave sources in the pulsar timing array waveband. For slowly evolving black hole binaries, their gravitational wave signals are continuous waves. For black hole encounters, their gravitational wave signals may appear as bursts. The following chapters describe data analysis methods for all of these sources.

1.3 Data analysis

The magnitude of pulse arrival time corrections induced by gravitational waves in the PTA waveband is $\sim 1 - 10$ns (Sesana et al., 2008, 2009). However, the rms of the stochastic timing noise contained in timing residuals for our current pulsar timing array is $\sim 30\text{ns} - 1.5\mu$s (Demorest et al., 2012). Therefore, the gravitational wave signals we try to detect by pulsar timing arrays are extremely weak compared with the timing noise, and correspondingly we need to do careful data analyses to achieve our goal. In this thesis, I mainly develope Bayesian analysis methods and apply them to analyze simulated pulsar timing data sets. Here I give a brief overview of Bayesian data analysis.

1.3.1 Overview of Bayesian data analysis

We denote $\boldsymbol{\theta}$ as the set of parameter values that characterizes the model of the pulse arrival time corrections induced by gravitational waves. To infer $\boldsymbol{\theta}$ is to estimate the probability density of $\boldsymbol{\theta}$ given the timing residuals $\mathbf{d}$ we have observed. This probability density is referred to as the posterior probability density, denoted $p(\boldsymbol{\theta}|\mathbf{d})$. Exploiting Bayes theorem, we can obtain:

$$p(\boldsymbol{\theta}|\mathbf{d}) = \frac{\Lambda(\mathbf{d}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{Z(\mathbf{d})} \tag{1.32}$$

where $\Lambda(\mathbf{d}|\boldsymbol{\theta})$ denotes the likelihood function representing the probability of observing timing residuals $\mathbf{d}$ given that the gravitational wave signal is characterized by the particular parameter set $\boldsymbol{\theta}$; $\pi(\boldsymbol{\theta})$ denotes our expectations, expressed as a probability density, of the values that $\boldsymbol{\theta}$ may take on prior to analyzing the data $\mathbf{d}$; and $Z(\mathbf{d})$ denotes normalization constant that makes $p(\boldsymbol{\theta}|\mathbf{d})$ a probability density.

The general framework for the Bayesian inference of $\boldsymbol{\theta}$ is to

(1) assign probability density to express our a priori expectation of $\boldsymbol{\theta}$;

(2) model the likelihood function of the timing residuals $\mathbf{d}$ we have observed;

(3) compute the posterior probability density via Eq. (1.32).

In a Bayesian analysis the posterior probability density contains the full inference of the parameter set $\boldsymbol{\theta}$.

1.3.2 Detection

To claim a detection of gravitational waves, we need to decide whether our observations do in fact include a gravitational wave signal. We adopt an information-theoretic approach to the problem of deciding whether we have detected the signal.

1.3.2.1 Introduction

Two different approaches to the problem of detection are found in the gravitational wave data analysis literature. The majority of the literature treats detection as a Neyman-Pearson hypothesis testing problem (e.g. Jenet et al., 2005, 2006; van Haasteren et al., 2009, 2011; Ellis et al., 2012b,c), while more recent work has approached the problem as one of Bayesian model selection (e.g. Finn & Lommen, 2010; Arzoumanian et al., 2014), to be decided via the calculation of the appropriate Bayes Factor (Kass & Raftery, 1995). Here we introduce a third approach, motivated by information theory, that makes use of the so-called Deviance Information Criterion (DIC, Spiegelhalter et al., 2002).

Why is a third approach needed? When we speak of detection the question we are interested in is how likely it is that we have observed a gravitational wave source. Neyman-Pearson hypothesis tests address a different question (Neyman & Pearson, 1933): assuming that a signal is absent, is the calculated value of the test statistic unusual (i.e., occurs in less than a fraction α of observations)? Notice that nothing about the signal that we are trying to detect enters into any calculations that we perform in evaluting the test statistic or its threshold and, not surprisingly, neither does the conclusion drawn address how much confidence we should ascribe to any claim of detection.

The Bayesian model comparison attempts to address the question of confidence by evaluating the data-dependent contribution to the relative probability of the two hypotheses "signal present" and "signal absent". This ratio of probabilities separates neatly into two terms. One of these, the so-called Bayes Factor, is strictly objective and depends only on the observations and the characteristics of the measurement noise. The other, which is data *independent,* depends only upon prior beliefs in the relative likelihood of the two hypotheses (Kass & Raftery, 1995). When there is a generally agreed upon prior belief in the relative probability of the two alternative hypotheses all observers will agree on exactly what an observation "says"; when there is disagreement it is always

traceable to disagreement in prior beliefs. In either case the Bayes Factor is an objective measure of what an observation contributes to the conclusion.

As appealing as the Bayes Factor is for assessing the data-derived evidence in favor of the hypothesis that a signal has been detected, in real applications it may suffer from shortcomings that render it useless or near useless. The Bayes Factor $\mathcal{B}$ associated with obserations $\mathbf{d}$ and the two alternative models $\mathcal{M}_0$ and $\mathcal{M}_1$ is the ratio of the probabilities of observing $\mathbf{d}$ under the different model assumptions (Kass & Raftery, 1995):

$$\mathcal{B} = \frac{\Lambda_{\mathcal{M}_1}(\mathbf{d}|\mathcal{M}_1)}{\Lambda_{\mathcal{M}_0}(\mathbf{d}|\mathcal{M}_0)}. \tag{1.33}$$

where $\Lambda(\mathbf{d}|\mathcal{M})$ is the probability of making the particular observation $\mathbf{d}$ under the hypothesis $\mathcal{M}$. Now suppose that $\mathcal{M}_0$ is the null hypothesis (i.e., $\mathbf{d}$ is noise only) and $\mathcal{M}_1$ is the hypothesis that a signal from, e.g., a gravitational wave, is present. The probability $\Lambda_{\mathcal{M}_1}(\mathbf{d}|\mathcal{M}_1)$ then depdends on the priors associated with the parameter vector $\boldsymbol{\theta}$ that characterizes a binary system in model $\mathcal{M}_1$:

$$\Lambda_{\mathcal{M}_1}(\mathbf{d}|\mathcal{M}_1) = \int \Lambda(\mathbf{d}|\boldsymbol{\theta}, \mathcal{M}_1)\pi(\boldsymbol{\theta}|\mathcal{M}_1)\,\mathrm{d}\boldsymbol{\theta}. \tag{1.34}$$

For an improper prior π the marginalized likelihood $\Lambda_{\mathcal{M}_1}(\mathbf{d}|\mathcal{M}_1)$, and consequently the Bayes Factor, may be indeterminate, even though the posterior probability density of the parameters may itself be proper. This high sensitivity of the Bayes Factor to the choice of prior is an example of the Lindley-Bartlett Paradox (Shafer, 1982).

We may try to transform the improper prior to a proper prior by imposing some upper or lower bounds on parameters and making the Bayes factor well defined. However, this procedure will not solve the problem either, because the consequent marginalized likelihood of $\mathcal{M}_1$, determined by Eq. (1.34), will depend on the cut-offs. As a result, the Bayes factor will be strongly sensitive to

the uncertain cut-offs and the evidence of gravitational waves will depend on the subjective choices of the upper or lower bounds of the parameters.

1.3.2.2 Deviance Information Criterion

The Deviance Information Criterion judges how well the data favors a model by the sum of two terms (Spiegelhalter et al., 2002) — one term represents "goodness of fitting", which measures how well the model fits the data and is the negative log likelihood of the model; the other term represents "the penalty of complexity", which measures the degree of overfitting or the effective number of parameters of the model (Spiegelhalter et al., 2002). The data favors the model with smaller DIC (Spiegelhalter et al., 2002).

To study the mathematical structure of DIC, we first review how the non-Bayesian analysis uses the combination of "goodness of fitting" and "penalty of complexity" to judge how well the data favors a model. In non-Bayesian analysis, people generally use the Akaike Information Criterion (AIC) to do model comparison (Akaike, 1974). The proposal of this criterion is motivated by the logic of Kullback-Leiber (KL) divergence (Kullback & Leibler, 1951). Suppose that the data set $\mathbf{y}$ is driven by a unknown process $\mathcal{S}$, and we use a model $\mathcal{M}$ to characterize $\mathcal{S}$. The KL divergence is defined as:

$$D_{KL}(\mathcal{S}|\mathcal{M}) = \int \mathcal{S} \log \frac{\mathcal{S}}{\mathcal{M}} \mathrm{d}\mathbf{y}. \tag{1.35}$$

This is basically the entropy difference of the process $\mathcal{S}$ and the model $\mathcal{M}$. It characterizes the information loss of using the model $\mathcal{M}$ to characterize the true process $\mathcal{S}$, i.e., the deviation of $\mathcal{M}$ from $\mathcal{S}$ (Kullback & Leibler, 1951). We would like to choose a model $\mathcal{M}$ that minimize the information loss. Akaike (1974) showed that for a large sample of data, the quantity

$$\mathrm{AIC} = 2k - 2\log(\hat{\Lambda}) \tag{1.36}$$

characterizes how much information is lost by using the model $\mathcal{M}$ to characterize $\mathcal{S}$, where k is the number of parameters used in the model and $\hat{\Lambda}$ is the maximum likelihood of the model. This quantity is called the Akaike Information Criterion. The second term $-2\log(\hat{\Lambda})$ shows how well the model fits the data ("goodness of fitting") and the parameter number k indicates the model complexity.

DIC is a Bayesian version of AIC. We define the deviance of a model as -2 times the log-likelihood (Kullback & Leibler, 1951):

$$D(\mathbf{d}, \boldsymbol{\theta}) = -2\log\Lambda(\mathbf{d}|\boldsymbol{\theta}) \tag{1.37}$$

If we set the parameter as the maximum likelihood estimator, then the deviance is same as the second term $-2\log(\hat{\Lambda})$ in Eq. (1.36). Therefore, this quantity characterizes the model discrepancy (Spiegelhalter et al., 2002) and so the average of the deviance on posterior probability distribution provides a summary of the error of the model $\mathcal{M}$ (Spiegelhalter et al., 2002):

$$D_{\text{avg}}(\mathbf{d}, \mathcal{M}) = \int D(\mathbf{d}, \boldsymbol{\theta}) p(\boldsymbol{\theta}|\mathbf{d}) \mathrm{d}\boldsymbol{\theta}. \tag{1.38}$$

With the deviance defined above, we can define the complexity of a model. In the model $\mathcal{M}$, the parameter set that represents the model is $\boldsymbol{\theta}$. In Bayesian analysis, we can use the mean value $\bar{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ under its posterior probability density to be the Bayesian estimator of $\boldsymbol{\theta}$ (Gelman et al., 2004, chap. 6),

$$\bar{\boldsymbol{\theta}} = \int \boldsymbol{\theta}\, p(\boldsymbol{\theta}|\mathbf{d})\, \mathrm{d}\boldsymbol{\theta}. \tag{1.39}$$

We can use the Bayesian estimator to represent the estimation of the model $\mathcal{M}$. Correspondingly, the excess of the true uncertainty over the estimated uncertainty will be denoted:

$$\Delta_D(\mathbf{d}, \boldsymbol{\theta}, \bar{\boldsymbol{\theta}}) = D(\mathbf{d}, \boldsymbol{\theta}) - D(\mathbf{d}, \bar{\boldsymbol{\theta}}) \tag{1.40}$$

which can be thought as the reduction in uncertainty due to Bayesian estimation, or alternatively the degree of overfitting due to $\bar{\boldsymbol{\theta}}$ adapting to the data set $\mathbf{d}$ (Spiegelhalter et al., 2002). Correspondingly, the mean value of Δ_D under the posterior probability density summarizes the model complexity or the degree of overfitting due to the model $\mathcal{M}$ adapting to the data set (Spiegelhalter et al., 2002; Meng & Rubin, 1992)

$$\begin{aligned} p_D(\mathbf{d}, \mathcal{M}_1) &= \int \Delta_D(\mathbf{d}, \boldsymbol{\theta}, \bar{\boldsymbol{\theta}})\, p(\boldsymbol{\theta}|\mathbf{d})\, \mathrm{d}\boldsymbol{\theta} \\ &= D_{\text{avg}}(\mathbf{d}, \mathcal{M}_1) - D(\mathbf{d}, \bar{\boldsymbol{\theta}}). \end{aligned} \tag{1.41}$$

We can see that when the posterior probability density of $\boldsymbol{\theta}$ is approximately a Gaussian distribution, p_D is approximately the trace of the product of the Fisher information matrix and the posterior covariance matrix, which is the effective number of parameters (Spiegelhalter et al., 2002). Therefore, p_D is similar to the model parameter number k in AIC (see Eq. 1.36). It is reasonable that the model complexity p_D depends on the observed data set because matching the data with the model would induce a statistical correlation among parameters that is likely to reduce the effective dimensionality of the model, and the degree of the reduction may depend on the specific data set (Spiegelhalter et al., 2002). Correspondingly the degree of overfitting of the model may depend on the data set.

In analogy to AIC in Eq. (1.36), evaluating the deviance Eq. (1.37) at the Bayesian estimator $\boldsymbol{\theta}$ characterizes the "goodness of fitting" (similar to the second term $-2\log(\hat{\Lambda})$ in Eq. (1.36)), and

p_D characterizes the model complexity. Following Eq. (1.36), Spiegelhalter et al. (2002) proposed a Bayesian version of AIC, i.e., Deviance Information Criterion:

$$\mathrm{DIC}(\mathbf{d}, \mathcal{M}) = D(\mathbf{d}, \bar{\boldsymbol{\theta}}, \mathcal{M}) + 2p_D(\mathbf{d}, \mathcal{M}). \tag{1.42}$$

This quantity is referred to as Deviance Information Criterion. The data would favor the model with smaller DIC, since such a model has a smaller discrepancy with the data and is less complex.

The difference between the DICs of two models $\mathcal{M}_1$ and $\mathcal{M}_0$,

$$\Delta\mathrm{DIC} = \mathrm{DIC}(\mathbf{d}, \mathcal{M}_1) - \mathrm{DIC}(\mathbf{d}, \mathcal{M}_0) \tag{1.43}$$

characterizes how much the timing residual observations favor $\mathcal{M}_1$ over $\mathcal{M}_0$. It is similar to twice the negative natural logarithm of the Bayes factor (Kass & Raftery, 1995). Correspondingly, it implies that $\exp(-\Delta\mathrm{DIC}/2)$ has the same scale as the Bayes factor (Spiegelhalter et al., 2002). If $\Delta\mathrm{DIC} \lesssim -10$, it implies that the equivalent Bayes factor between the two models is ~ 150. So it is safe to conclude that the data strongly favors $\mathcal{M}_1$ and there is strong evidence that a gravitational wave is present in the data set (Spiegelhalter et al., 2002).

1.3.2.3 Comparison of DIC and Bayes factor

Unlike the Bayes factor, the DIC is sensitive to the posterior probability distribution, which must be proper, but not the prior probability distribution, which could be improper. Following Eq. (1.32), the posterior distribution is the product of the data-dependent likelihood function and the data-independent prior distribution. A robust estimation of the gravitational wave parameter set should be achieved primarily by the data-dependent likelihood function but not the data-independent prior expectation. Correspondingly, when the evidence for gravitational waves is

strong, the posterior distribution is determined primarily by the likelihood function and is insensitive to the prior distribution. Therefore, the DIC is insensitive to the prior distribution, even if the prior distribution is improper or has highly uncertain cutoffs.

The main weakness of DIC is that it is a quantity motivated by the AIC and KL divergence, but not a rigorously derived one from statistics principles. In particular, the definition of the model complexity p_D is ad hoc. However, given the premature status of the research on Bayesian model comparison (Spiegelhalter et al., 2002), DIC may be one of the best choices.

1.3.3 Computational method

We need to robustly compute and sample the posterior probability distribution to estimate the gravitational wave parameter set and infer whether the evidence for gravitational waves is strong. When the gravitational wave parameter set is high dimensional, the simulation of the posterior probability distribution requires a highly efficient computational method. Markov Chain Monte Carlo (MCMC) is designed for this purpose (Robert & Casella, 2004) and I will give a brief summary here.

1.3.3.1 Markov Chain Monte Carlo

The Markov Chain Monte Carlo method is a stochastic method to sample a high dimensional probability distribution. The Metropolis-Hastings algorithm is one of the MCMC methods (Metroplis et al., 1953; Hastings, 1970) that I primarily use to simulate the posterior probability distribution in the case of gravitational wave simulated data analysis.

To simulate the posterior probability distribution $p(\boldsymbol{\theta}|\mathbf{d})$, the Metropolis-Hastings algorithm follows the following procedure:

(1) Choose an initial parameter set $\boldsymbol{\theta}_0$;

(2) Choose a proposal distribution q conditioned on $\boldsymbol{\theta}_0$, and sample a parameter set $\tilde{\boldsymbol{\theta}}$ from the conditional distribution q, i.e.,

$$\tilde{\boldsymbol{\theta}} \sim q(\cdot|\boldsymbol{\theta}_0) \tag{1.44}$$

(3) Take the next step parameter set $\boldsymbol{\theta}_1$ as

$$\boldsymbol{\theta}_1 = \begin{cases} \tilde{\boldsymbol{\theta}}, & \text{with probability} \quad \rho(\boldsymbol{\theta}_0, \tilde{\boldsymbol{\theta}}) \\ \boldsymbol{\theta}, & \text{with probability} \quad 1 - \rho(\boldsymbol{\theta}_0, \tilde{\boldsymbol{\theta}}) \end{cases} \tag{1.45}$$

where

$$\rho(\boldsymbol{\theta}_0, \tilde{\boldsymbol{\theta}}) = \min \left\{ \frac{p(\tilde{\boldsymbol{\theta}}|\mathbf{d})}{p(\boldsymbol{\theta}_0|\mathbf{d})} \frac{q(\boldsymbol{\theta}_0|\tilde{\boldsymbol{\theta}})}{q(\tilde{\boldsymbol{\theta}}|\boldsymbol{\theta}_0)} \right\} \tag{1.46}$$

(4) repeat step (1) (2) (3) and generate a chain of parameter set $\boldsymbol{\theta}$. Such a chain of $\boldsymbol{\theta}$ forms a sample of the posterior distribution $p(\boldsymbol{\theta}|\mathbf{d})$ (Robert & Casella, 2004, chap. 7).

1.3.3.2 Convergence diagnostics

We must obtain a convergent sample of the parameter set $\boldsymbol{\theta}$ in order to robustly simulate the posterior probability distribution $p(\boldsymbol{\theta}|\mathbf{d})$. To achieve this goal, the sample we draw by the Metropolis-Hastings algorithm must be of sufficient size with appropriate burning in and thinning steps [1]. In order to improve the convergence of the MCMC sample, we usually draw several chains independently with different initial conditions and then combine all the chains, which further reduces the effect of initial values and autocorrelation (Robert & Casella, 2004, chap. 12).

[1] Burning in M steps means throwing away the first M steps of the MCMC chain of $\boldsymbol{\theta}$ to minimize the effect of initial values. Thinning k steps means keeping only every kth step of the chain to reduce any autocorrelation of the sample (Robert & Casella, 2004, chap. 12).

There are several methods to check the convergence of the MCMC sample (Robert & Casella, 2004, chap. 12) and I use the spectral analysis proposed by Geweke (1992) in my thesis. The method is to check the similarity between the first part and the last part of the MCMC sample. Assuming that the MCMC sample contains N, Geweke (1992) proposes the Z-score of the sample as

$$Z = \frac{\mu_A - \mu_B}{\sqrt{\frac{\sigma_A^2}{N_A} + \frac{\sigma_B^2}{N_B}}} \tag{1.47}$$

where μ_A and μ_B are respectively the mean values of the first N_A and the last N_B draws of the sample; and σ_A^2 and σ_B^2 are respectively the variances of the first N_A and the last N_B draws of the sample. The values of N_A and N_B suggested by Geweke (1992) are $N_A/N = 0.1$ and $N_B/N = 0.5$. When the MCMC sample is multi-dimensional, we need to compute the Z-score for each of the parameters.

In the following chapters, I use the Metropolis-Hastings algorithm to simulate posterior probability distributions of all the parameters in all different simulated pulsar timing data sets. For each simulated data set, I sample 20 chains with 5 million draws for each chain. I burn-in 1 million draws and thin every 100-step for each chain. I will show that the absolute values of the Z-scores are all less than 1/3, which indicates the probability that the samples are convergent is greater than 74%.

1.4 Application of gravitational wave data analysis methods to detect exoplanetary systems

1.4.1 Overview of exoplanet detection

Over the last decade, improvements in astronomical instrumentation have led to rapid progress in the detection and study of exoplanets and exoplanetery systems. There are primarily two techniques that lead to the discovery of exoplanets (Roques et al., 2014):

(1) Radial velocity — use Doppler spectroscopy to measure the radial velocity variations [2] of the stars that may be induced by the gravity of the orbiting planets (Lovis & Fischer, 2010). So far there have been about 592 exoplanets detected by this method (Roques et al., 2014).

(2) Transit photometry — use a photometric method to measure the small drop of the brightness of the stars due to the transits of surrounding planets in front of the stars (Winn, 2010). So far there have been about 1189 exoplanets detected by this method thanks to the launch of the Kepler Space Telescope (Roques et al., 2014).

There are also several other techniques that lead to the discovery of a small fraction of planets such as gravitational microlensing, direct imaging, pulsar timing, astrometry, etc. (Seager, 2010; Perryman, 2014).

1.4.2 Radial velocity planet detection

The radial velocity (RV) method detects planets indirectly, observing the velocity variations of the host star along the line of sight that are induced by the motion of the host star about the star-planet system's center of mass. These radial velocity variations can be detected spectroscopically,

[2] Radial velocity is the velocity of the star in the direction of the line of sight between the star and the Earth

observing the small Doppler shifts of the star's spectral lines due to its motion about the system barycenter (Lovis & Fischer, 2010).

Assuming that the orbit of a planet of mass m_2 around a host star of mass m_1 is a Keplerian orbit with a period P and an eccentricity e, the radial velocity variation of the host star induced by the surrounding planet is then (Lovis & Fischer, 2010)

$$V_r = K\left[\cos(\Theta + \omega) + e\cos\omega\right] \tag{1.48a}$$

where ω is the longitude of periastron passage; Θ is the true anomaly of the elliptical Keplerian orbit; and K is the semiamplitude of the RV variations.

The true anomaly Θ is the solution of Kepler's equation,

$$\tan\left(\frac{\Theta}{2}\right) = \left(\frac{1+e}{1-e}\right)^{1/2}\tan\left(\frac{E}{2}\right) \tag{1.48b}$$

$$E - e\sin(E) = \frac{2\pi}{P}(t - t_0) \tag{1.48c}$$

where E mean anomaly and t_0 is the time of periastron passage.

The semiamplitude of the RV variations K induced by the surrounding planet can be expressed as

$$K = 28.43\,\mathrm{m\,s^{-1}}\left(\frac{1}{\sqrt{1-e^2}}\right)\left(\frac{m_2\sin\iota}{\mathrm{M_{Jupiter}}}\right)\left(\frac{m_1+m_2}{\mathrm{M_\odot}}\right)^{-2/3}\left(\frac{P}{\mathrm{yr}}\right)^{-1/3} \tag{1.49}$$

where ι is the orbital inclination angle.

Therefore, based on Eq. (1.48), if we detect a planet in RV data, we can determine its five Keplerian orbital parameters: semiamplitude K, orbital period P, orbital eccentricity e, longitude of periastron passage ω and time of periastron passage t_0.

1.4.3 Detecting exoplanets around giant stars

Although the majority of planets detected to date have focused on dwarfs (Roques et al., 2014), planets have also been found around giant stars (e.g. Sato et al., 2003; Gettel, 2012; Niedzielski et al., 2015). It is important to study planetary systems around a broad range of evolutionary states of host stars, to ensure that we have a complete picture of planet formation and evolution. In main sequence A-F stars with masses $\gtrsim 1.5 M_{\odot}$, the lower number of spectral features and their rotational broadening make it difficult to reach a high RV precision of Doppler spectroscopy (Lovis & Fischer, 2010; Gettel, 2012). As a result, we have little direct information about the formation and evolution of planets around intermediate mass stars. As these stars evolve off the main sequence and into giant stars, their atmospheres cool and form many narrow spectral lines, making them accessible to radial velocity planet searches (Gettel, 2012). Correspondingly, detecting planets around giant stars is particularly valuable, as it allows us to study the planetary systems around intermediate mass stars.

The first robust detection of a planet around a giant star involved HD 104985 (Sato et al., 2003) from a survey of 300 stars at the Okayama Astrophysical Observatory. So far there have been about 50 planets around giant stars detected (Niedzielski et al., 2015). The Penn State-Torun Planet Search (PTPS), using the Hobby Eberly Telescope (HET), is currently the largest survey of planets around giant stars with $\sim$ 1000 targets (Niedzielski et al., 2007; Zieliński et al., 2012; Adamów et al., 2014). My work on detecting exoplanets around giant stars is in collaboration with the PTPS group.

1.5 Outline

Chapter 2 proposes a method to determine the luminosity distances of nearby gravitational wave sources by measuring the gravitational wave phasefront curvature via a pulsar timing array.

This work was published in the Monthly Notices of Royal Astronomical Society 414, 58 (2011). Chapter 3 proposes a Bayesian nonparametric method aimed at detecting gravitational wave bursts from PTA data. This work was published in Physical Review D 90, 104029 (2014). Chapter 4 proposes a Bayesian nonparametric method aimed at detecting a generic gravitational wave background via PTAs. This work was published in Physical Review D 90, 024020 (2014). Chapter 5 proposes a Bayesian parametric method aimed at detecting gravitational waves from slowly evolving supermassive black hole binaries via PTAs. In Chapter 6, I describe the application of Bayesian analysis to analyze a set of radial velocity data collected by the Hobby-Eberly Telescope to detect an exoplanetary system around a K0 giant star.

Chapter 2

Pulsar Timing Array Observations of Gravitational Wave Source Timing Parallax

2.1 Introduction

As a gravitational wave detector a PTA is large: its size L (on order of kpc) is much greater than the gravitational radiation wavelength scale λ (on order of pc) that we use to probe it. The conventional approximation of gravitational waves as plane-fronted applies only for sources at distances $R \gg L^2/\lambda$. Correspondingly, for point gravitational wave sources closer than $\sim$ 100Mpc, the curvature of the gravitational radiation phasefronts contributes significantly to the detector response. Analyses focused on the discovery of such sources must take proper account of the curvature of the radiation phasefront. In this chapter we evaluate the PTA response to gravitational wave point sources including the important wavefront curvature effects. Taking the wavefront curvature into account the relative amplitude and phase of the timing residuals associated with a collection of pulsars allow us to measure the distance to, and sky position of, the source.

In §2.2 we evaluate the pulse arrival time disturbance associated with a spherically-fronted gravitational wave traversing a pulsar-Earth line-of-sight and compare it with the response to the same wave in the plane-wave approximation. In §2.3 we describe how, owing to their sensitivity to gravitational radiation phasefront curvature, PTA observations of point sources can be used to measure or place lower bounds on the distance to the source and, more surprisingly, the distance to the array pulsars. Finally we summarize our conclusions in §2.4.

This work was published as X. Deng & L. S. Finn, 2011, Monthly Notices of Royal Astronomical Society 414, 58.

2.2 Pulsar timing residuals from spherically-fronted gravitational waves

2.2.1 Timing residuals

We focus our attention on the electromagnetic field associated with the pulsed emission of a pulsar and denote the field phase, at the pulsar, as $\phi_0(t)$. We are interested in the time-dependent phase $\phi_0(t)$ of the electromagnetic field associated with the pulsed emission measured at an Earth-based radio telescope, which we write as

$$\phi(t) = \phi_0[t - L - \tau_0(t) - \tau_{\rm GW}(t)] \tag{2.1a}$$

where

$$\tau_0 = \left(\begin{array}{l} \text{Corrections owing exclusively to the spatial motion of the Earth} \\ \text{within the solar system, the solar system with respect to the pulsar,} \\ \text{and electromagnetic wave propagation the interstellar medium} \end{array} \right) \tag{2.1b}$$

$$\tau_{\rm GW} = \left(\text{ Corrections owing exclusively to a gravitational wave } \right) \tag{2.1c}$$

$$L = (\text{Earth-pulsar distance})\,. \tag{2.1d}$$

(Note that we work in units where $c = G = 1$.)

In the absence of gravitational waves $\tau_{\rm GW}$ vanishes and the phase front $\phi_0(t)$ arrives at Earth at time $t_\oplus(t) = t + L + \tau_0(t)$. In the presence of a gravitational wave signal the phase front arrives at time $t_\oplus(t) + \tau_{\rm GW}(t)$; thus, $\tau_{\rm GW}$ is the gravitational wave timing residual. Following Finn (2009) Eqs. (3.26) and (3.12e), the arrival time correction $\tau_{\rm GW}(t)$ is

$$\tau_{\rm GW}(t) = -\frac{1}{2}\hat{n}^l\hat{n}^m\mathcal{H}_{lm}(t) \tag{2.2a}$$

where $\mathcal{H}_{lm}$ is the integral of the transverse-traceless metric perturbation over the null geodesic ranging from the pulsar to Earth:

$$\mathcal{H}_{lm}(t) = L \int_{-1}^{0} h_{lm}\left(t + L\xi, \mathcal{P} - L(1+\xi)\hat{n}\right)\, d\xi \tag{2.2b}$$

$$\mathcal{P} = (\text{Pulsar location}) \tag{2.2c}$$

$$\hat{n} = (\text{Unit vector pointing from Earth to pulsar})\,. \tag{2.2d}$$

2.2.2 Gravitational waves from a compact source

At a point located at $\mathcal{X}$ in the perturbative regime far from the source the transverse-traceless gauge gravitational wave metric perturbation associated with radiation from a compact source may be written (Finn, 1985)

$$h_{lm}(t, \mathcal{X}) = \int_{-\infty}^{\infty} \widetilde{h}_{lm}(f, \hat{k}) e^{-2\pi i f t} df \tag{2.3a}$$

$$= \frac{1}{|\mathcal{X} - \mathcal{S}|} \int_{-\infty}^{\infty} \left[\widetilde{A}_{lm}(f, \hat{k}) e^{2\pi i f |\mathcal{X}-\mathcal{S}|}\right] e^{-2\pi i f t} df \tag{2.3b}$$

where

$$\mathcal{S} = (\text{source location}) \tag{2.3c}$$

$$\hat{k} = \frac{\mathcal{X} - \mathcal{S}}{|\mathcal{X} - \mathcal{S}|} = (\text{unit vector in direction of wave propagation at } \mathcal{X})\,. \tag{2.3d}$$

The Fourier coefficient functions $\widetilde{A}_{lm}(f, \hat{k})$ are everywhere symmetric, traceless and transverse with respect to $\mathcal{X} - \mathcal{S}$: i.e.,

$$\widetilde{A}_{lm}(f, \hat{k}) = \widetilde{A}_{+}(f, \hat{k}) \mathbf{e}^{(+)}_{lm} + \widetilde{A}_{\times}(f, \hat{k}) \mathbf{e}^{(\times)}_{lm} \tag{2.4}$$

where $\mathbf{e}^{(+)}_{lm}(\hat{k})$ and $\mathbf{e}^{(\times)}_{lm}(\hat{k})$ are the usual transverse-traceless gravitational wave polarization tensors for waves traveling in direction $\hat{k}$.

2.2.3 Response function

With expression 2.3 for h_{lm} we can evaluate $\mathcal{H}_{lm}$ in the Fourier domain:

$$\widetilde{\mathcal{H}}_{lm}(f) = L \int_{-1}^{0} \widetilde{h}_{lm}\left(f, \mathcal{P} - L(1+\xi)\hat{n}\right) e^{-2\pi i f L \xi}\, d\xi \tag{2.5a}$$

$$= L \int_{-1}^{0} \frac{\widetilde{A}_{lm}\left(f, \hat{k}(\xi)\right)}{|\mathcal{P} - L(1+\xi)\hat{n} - \mathcal{S}|} \exp\left[2\pi i f \left(-L\xi + |\mathcal{P} - L(1+\xi)\hat{n} - \mathcal{S}|\right)\right] d\xi \tag{2.5b}$$

Figure 2.1 describes the Earth-pulsar-source geometry: note we define R to be the Earth-source distance and θ the angle between the Earth-pulsar and Earth-source lines-of-sight. Along the path traversed by the pulsar's electromagnetic pulse phase fronts ($\mathcal{P} - L\xi\hat{n}$ for $-1 \leq \xi \leq 0$) we may write

$$\hat{k}(\xi) = \frac{\mathcal{P} - L(1+\xi)\hat{n} - \mathcal{S}}{|\mathcal{P} - L(1+\xi)\hat{n} - \mathcal{S}|} \tag{2.6a}$$

$$r(\xi) = (\text{distance between gravitational wave phase front and source})$$

$$= |\mathcal{P} - L(1+\xi)\hat{n} - \mathcal{S}| = R\left[1 + 2\frac{L\xi}{R}\cos\theta + \left(\frac{L\xi}{R}\right)^2\right]^{1/2}. \tag{2.6b}$$

From Equations 2.5 and 2.6b we note that the time of arrival disturbance owing to a passing gravitational wave depends on the dimensionless quantities $\pi f L$ and L/R. PTA pulsar distances L are all of order kpc. Gravitational waves detectable via pulsar timing observations have frequencies greater than the inverse duration of the observational data set and less than the typical sampling period: i.e., $10^{-5}\ Hz \gtrsim f \gtrsim 10^{-9}$ Hz. Strong sources of gravitational waves (e.g., supermassive black hole binaries and triplets) in this band are all expected to be extragalactic, with significant

numbers within roughly 100 Mpc (Sesana & Vecchio, 2010). The scales of interest are thus either very large or very small

$$\epsilon = 10^{-5} \left(\frac{L}{1\,\mathrm{kpc}}\right) \left(\frac{100\,\mathrm{Mpc}}{R}\right) \tag{2.7a}$$

$$\pi f L = 10^{4} \left(\frac{f}{1\,\mathrm{yr}^{-1}}\right) \left(\frac{L}{1\,\mathrm{kpc}}\right) \tag{2.7b}$$

$$\pi f L \epsilon = 10^{-1} \left(\frac{f}{1\,\mathrm{yr}^{-1}}\right) \left(\frac{L}{1\,\mathrm{kpc}}\right)^{2} \left(\frac{100\,\mathrm{Mpc}}{R}\right), \tag{2.7c}$$

where we have introduced ϵ to denote the dimensionless L/R. Taking advantage of these scales the Fourier transform $\widetilde{\mathcal{H}}_{lm}(f)$ may be expressed

$$\widetilde{\mathcal{H}}_{lm}(f) = L \int_{-1}^{0} \widetilde{h}_{lm}\left(f, \mathcal{P} - L(1+\xi)\hat{n}\right) e^{-2\pi i f L \xi}\, d\xi \tag{2.8a}$$

$$= L \int_{-1}^{0} \frac{\widetilde{A}_{lm}\left(f, \hat{k}(\xi)\right)}{|\mathcal{P} - L(1+\xi)\hat{n} - \mathcal{S}|} e^{2\pi i f(-L\xi + |\mathcal{P} - L(1+\xi)\hat{n} - \mathcal{S}|)}\, d\xi \tag{2.8b}$$

$$= \widetilde{A}_{lm}(f, \hat{k}_0) \frac{\exp\left[\pi i f R \left(2 - \frac{1-\cos\theta}{1+\cos\theta}\right)\right]}{\sqrt{fR}\sin\theta} \frac{e^{-i\pi/4}}{2} \operatorname{erf}\left[e^{i\pi/4} x\right]\Big|_{x=\frac{\sqrt{\pi f R}\sin\theta}{1+\cos\theta}}^{\sqrt{\pi f R}\frac{\epsilon(1+\cos\theta)+1}{1+\cos\theta}\sin\theta}$$

$$\times \left[1 + \mathcal{O}(\epsilon) + \mathcal{O}\left(\epsilon \frac{\partial \log \tilde{A}}{\partial \hat{k}} \cdot \hat{n}\right)\right]. \tag{2.8c}$$

Except when $\theta \leq (\pi f R)^{-1/2} \ll 1$ the error function arguments are always large in magnitude. Taking advantage of the asymptotic expansion of $\operatorname{erf}(z)$ about the point at infinity,

$$\operatorname{erf}(z) \sim 1 - \frac{\exp\left(-z^2\right)}{z\sqrt{\pi}} \tag{2.9}$$

we may thus write

$$\widetilde{\tau}_{\rm gw} = -\frac{1}{2}\hat{n}^{l}\hat{n}^{m}\widetilde{\mathcal{H}}_{lm}(f) \tag{2.10a}$$

$$= \widetilde{\tau}_{\rm pw} + \widetilde{\tau}_{\rm cr} \tag{2.10b}$$

where

$$\widetilde{\tau}_{\rm pw} = \text{(plane wave approximation timing residual)}$$

$$= 2\widetilde{\mathcal{A}}(f,\hat{k}_0)\exp\left(2\pi i f L\sin^2\frac{\theta}{2}\right)\mathrm{sinc}\left(2\pi f L\sin^2\frac{\theta}{2}\right) \tag{2.10c}$$

$$\widetilde{\tau}_{\rm cr} = \text{(phasefront curvature correction to plane wave timing residual)}$$

$$= \epsilon\,(1+\cos\theta)\,\widetilde{\mathcal{A}}(f,\hat{k}_0)\exp\left[\pi i f L\left(4\sin^2\frac{\theta}{2}+\frac{\epsilon}{2}\sin^2\theta\right)\right]\mathrm{sinc}\left[\frac{\pi f L\epsilon}{2}\sin^2\theta\right] \tag{2.10d}$$

$$\widetilde{\mathcal{A}}(f,\hat{k}_0) = \frac{i\epsilon}{4}\left[\hat{n}^{l}\hat{n}^{m}\widetilde{A}_{lm}(f,\hat{k}_0)\right]\exp\left[2\pi i f R\right] \tag{2.10e}$$

and

$$\mathrm{sinc}(x) = \left(\begin{array}{c}\text{Unnormalized}\\ \\ \text{sinc function}\end{array}\right) = \frac{\sin(x)}{x}. \tag{2.10f}$$

These expressions are valid for all $\pi f L \gg 1$.

2.3 Discussion

The Fourier amplitude and phase of the gravitational wave timing residual $\widetilde{\tau}_{\rm GW}$ for any particular pulsar depend on the pulsar's position and distance relative to the position and distance of the gravitational wave source, and the relative amplitude of the signal in the two gravitational wave

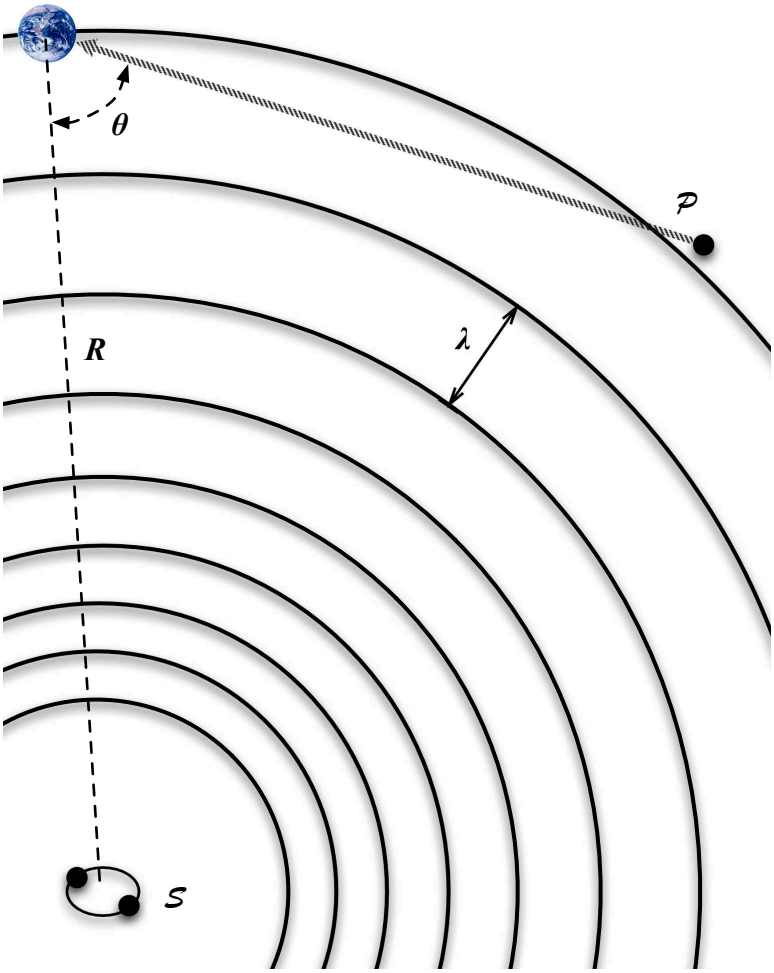


Fig. 2.1 Gravitational wave source (depicted as a binary system), Earth and pulsar. The gravitational wave phasefronts intersect the pulsar-Earth path, traveled by the electromagnetic pulses, in a manner that depends on the distance to the source and the source-Earth-pulsar angle θ.

polarizations. Correspondingly, observation in four or more pulsars of the gravitational wave timing residuals associated with a single source can, in principle, be used to infer the three-dimensional source position and location on the sky. In this section we discuss the conditions under which sufficiently accurate measurements of the $\widetilde{\tau}_{\rm GW}$ are possible and provide a proof-of-principle demonstration of how they may be combined to determine the sky location and distance to a gravitational wave source.

2.3.1 Plane wave approximation timing residuals

We focus our attention first on $\widetilde{\tau}_{\rm pw}$: the plane-wave approximation to $\widetilde{\tau}_{\rm gw}$. In order to relate the angular location of a gravitational wave source to the relative amplitude and phase of the $\widetilde{\tau}_{\rm pw}$ for two or more PTA pulsars the uncertainties $\sigma_{_L}$ in the Earth-pulsar distances, and σ_θ in pulsar sky locations, must satisfy

$$\sigma_{_L} \lesssim 2\,{\rm pc}\left(\frac{0.1\,{\rm yr}^{-1}}{f}\right)\left(\frac{1/2}{\sin^2(\theta/2)}\right) \tag{2.11a}$$

$$\sigma_\theta \lesssim 2.5'\left(\frac{0.1\,{\rm yr}^{-1}}{f}\right)\left(\frac{\rm kpc}{L}\right)\left(\frac{\pi/4}{|\sin\theta|}\right). \tag{2.11b}$$

The constraint on pulsar angular position accuracy is not a severe one; however, the constraint on pulsar distance accuracy is quite significant. For this reason previous pulsar timing searches for periodic gravitational waves (Yardley et al., 2010; Jenet et al., 2004; Lommen & Backer, 2001; Lommen, 2001) have sought only to identify an anomalous periodic contribution to pulse arrival times, avoiding explicit use of the amplitude of $\widetilde{\tau}_{\rm gw}$.

Improvements in instrumentation and timing techniques are making possible pulsar distance measurements of the accuracy and precision necessary to relate the $\widetilde{\tau}_{\rm pw}$ to the gravitational wave source angular location. For example, in recent years a succession of VLBI observations have

established the distance to J0437 − 4715 with an accuracy of ±1.3 pc (Verbiest et al., 2008; Deller et al., 2008; Deller, 2009); correspondingly, for J0437 − 4715 predictions of the amplitude and phase of $\widetilde{\tau}_{\rm pw}$ can be made and take part in the analysis of timing data to search for gravitational waves at frequencies $f \lesssim 0.1\,{\rm yr}$. Accurate pulsar distance measurements have been identified as a crucial element of a continuing program to use precision pulsar timing and interferometry to test gravity and measure neutron star properties; correspondingly, we can reasonably expect that the number of pulsars with high-precision distances will increase rapidly over the next several years.

Looking toward the future, observations with the SKA (Smits et al., 2009) using existing timing techniques are expected to be capable of measuring timing parallax distances to millisecond pulsars at 20 kpc with an accuracy of 20% (Smits et al., 2011). When measured by VLBI or timing parallax the fractional uncertainty in the pulsar distance is equal to the fractional uncertainty in the semi-annual pulse arrival time variation owing to the curvature of the pulsar's electromagnetic phasefronts: i.e.,

$$\frac{\sigma_L}{L} = \frac{\sigma_\tau}{\Delta\tau} \tag{2.12a}$$

where σ_τ^2 is the timing noise power in the bandwidth T^{-1} (for T the observation duration) about a frequency of $2/yr$ and

$$\Delta\tau = \left(\frac{\rm au}{2c}\right)\left(\frac{\rm au}{L}\right) = 1.2\,\mu{\rm s}\,\left(\frac{\rm kpc}{L}\right) \tag{2.12b}$$

is the peak-to-peak variation in the pulse arrival time residual owing to the curvature of the pulsar's electromagnetic phasefronts. Correspondingly, a 20% accuracy measurement of the distance to a pulsar at 20 kpc corresponds to a 0.5% accuracy measurement of the distance to a pulsar of the σ_τ at 500 pc: i.e., 2.5 pc.

This estimate of $\sigma_L = 2.5$ pc for a pulsar at 500 pc assumes constant signal-to-noise and, correspondingly, constant timing precision σ_τ. Closer pulsars will have greater fluxes and, correspondingly, better timing precision; consequently we may regard this as an upper bound on the distance measurement error σ_L. If we assume that σ_τ^{-2} is proportional to the pulsar energy flux then

$$\sigma_\tau = 0.6\,\mathrm{ns}\,\left(\frac{L}{\mathrm{kpc}}\right) \tag{2.13a}$$

$$\sigma_L = 0.5\,\mathrm{pc}\,\left(\frac{L}{\mathrm{kpc}}\right)^3; \tag{2.13b}$$

i.e., if the only barrier to improved timing precision is signal-to-noise then sub-parsec precision distance measurements should be possible in the SKA era for pulsars within a kpc.[1]

In practice other contributions to the noise budget will eventually limit the timing precision. Among these the most important are intrinsic pulsar timing noise, pulse broadening owing to scattering in the interstellar medium (ISM) and tropospheric scintillation. Recent improvements in pulsar timing techniques (Lyne et al., 2010) suggest that intrinsic pulsar timing noise need not limit timing precision measurements in either present or future instrumentation. Unresolved pulse broadening owing to scattering in the ISM will limit the attainable σ_τ and, thus, σ_L. For nearby ($L < 1$ kpc) pulsars the *measured* pulse broadening at $\nu = 1$ GHz is typically in the 1–10 ns range (Manchester et al., 2005). Since pulse broadening scales with observation frequency as $\nu^{-4.4}$ (though observations suggest the index may be closer to -3.9 (Bhat et al., 2004)) scattering in the ISM also need not limit our ability to measure the distance to nearby pulsars. Finally,

[1] These estimates are confirmed by recent work of Lee et al. (2011), who consider the role played by pulsar timing parallax in identifying gravitational waves from supermassive black hole binaries. Corbin & Cornish (2010) have also looked at how well pulsar distances may be determined from gravitational wave observations of "chirping" black hole binary systems, where the combination of period evolution, source location, and signal amplitude in the gravitational wave polarizations is sufficient to determine the source distance. They find modest improvements in array pulsar distance estimates over a priori estimates may be possible, but are by no means certain.

recent estimates by Jenet et al. (2011) indicate that, for gravitational wave frequencies $\lesssim \mathrm{yr}^{-1}$, tropospheric scintillation will contribute to the noise budget at no more than the 0.03 ns rms level. Consequently, over the next decade $\sigma_L \lesssim 1\,\mathrm{pc}$ is thus a reasonable expectation for millisecond pulsars at distances up to a kpc.

2.3.2 Correction to plane wave approximation

We turn now to the ratio of Fourier amplitudes

$$\begin{aligned}\rho(f,\hat{k}_0) &= \frac{\widetilde{\tau}_{\mathrm{cr}}}{\widetilde{\tau}_{\mathrm{pw}}} \\ &= \exp\left[\pi i f L\left(2\sin^2\frac{\theta}{2}+\frac{\epsilon}{2}\sin^2\theta\right)\right] \\ &\quad\times\frac{\sin\left(\frac{1}{2}\pi f L\epsilon\sin^2\theta\right)}{\sin\left(2\pi f L\sin^2\frac{\theta}{2}\right)}.\end{aligned} \tag{2.14}$$

When $\pi f L\epsilon\sin^2\theta \geq 1$ each of the sin functions whose ratio determines the ratio $|\rho|$ has order-unity magnitude and $|\rho|$ oscillates rapidly from 0 to ∞. For a typical PTA pulsar (distance $L \simeq$ kpc) this will be the case for source frequencies greater than or of order yr^{-1} and distances R less than or of order 100 Mpc (cf. Eqs. 2.7). In this regime the gravitational wave phasefront curvature plays an essential role in determining the response of a pulsar timing array.

As an example consider the residuals in the pulse arrival times for PSR J1939 + 2134 owing to gravitational waves from a supermassive black hole binary system like that proposed by Sudou et al. (2003) in the radio galaxy 3C66B at RA 2h23m11.4s, Dec 42°59'31" and luminosity distance $R = 85.8$ Mpc (Sudou et al., 2003). The binary, ruled-out by the subsequent analysis of Jenet et al. (2004), was supposed to have a total mass of $5.4\times10^{10}\mathrm{M}_\odot$, a mass ratio of 0.1, and a period of 1.05 yr, corresponding to a gravitational wave frequency of $1.9\,\mathrm{yr}^{-1}$. Assuming these binary parameters to be exact and taking the distance and location of J1939 + 2134 as provided by the

ATNF pulsar catalog (Manchester et al., 2005) yields $\rho = 1.1\exp[-0.86\pi i]$: i.e., the magnitude of the "correction" $\tau_{\rm cr}$ is greater than the "leading-order" plane-wave approximation contribution, the magnitude of $\tau_{\rm gw}$ is 46% of that predicted by the plane wave approximation, and the phase of $\tau_{\rm gw}$ is retarded by $\pi/2$ radians relative to $\tau_{\rm pw}$.[2]

2.3.3 Gravitational wave source distance and location on sky

The magnitude and phase of the plane-wave approximation Fourier coefficient function $\widetilde{\tau}_{\rm pw}$ depends on the distance to the associated pulsar and the pulsar-Earth-source angle. The correction depends on these and, in addition, the source distance. Measurements of $\widetilde{\tau}_{\rm gw}$ in a collection of pulsars whose angular location and distance are known sufficiently accurately can thus be used to measure the curvature of the gravitational radiation phase-fronts and, thus, the luminosity distance and direction to the gravitational wave source.

Consider a periodic source of gravitational waves observed in an array of pulsars whose distances and relative locations are known to high precision. The gravitational wave contribution to the pulse arrival times for array pulsar ℓ may be represented by an amplitude and phase, which will depend on the different $\epsilon_\ell = L_\ell/R$, $\pi f L_\ell$ and θ for each pulsar ℓ in the array. Requiring that these phases and amplitudes all be consistent is a powerful constraint on the source distance and location on the celestial sphere.

As a crude but effective proof-of-principle demonstration of distance measurement by gravitational wave timing parallax consider observations made with a selection of International Pulsar Timing Array (IPTA) pulsars with distances whose precision has been projected, as described above, into the SKA era. We focus our attention on the example source in 3C66B. For this source

[2]This comparison is meant only to illustrate the importance of $\tau_{\rm cr}$ in estimating $\tau_{\rm gw}$. In this case the distance to J1939 + 2134 is not known accurately enough to allow a prediction for an accurate estimate of either $\tau_{\rm pw}$ or $\tau_{\rm gw}$.

$(\pi f)^{-1} = 0.051\,\mathrm{pc}$. Following Equation 2.13, four pulsars currently monitored as part of the International Pulsar Timing Array are close enough that we may expect their distance to be measured with a one-sigma uncertainty of less than 0.025 pc. Table 2.1 lists these pulsars and their currently measured distances. For these four pulsars:

1. Calculate the four $\widetilde{\tau}_{\mathrm{gw}}$ for the example source in 3C66B described above. These complex amplitudes, corresponding to periodic timing residual amplitude and phase, are our "observations". Denote the $\widetilde{\tau}_{\mathrm{gw}}$ for pulsar k by $\widetilde{\tau}_k$.

2. Adopt approximate distances to each of these four pulsars consistent with normally distributed measurement error with standard deviation as given in Equation 2.13.

3. Using these approximate distances evaluate

$$\psi^2(r) = \sum_k \left| \log\left(\frac{\widetilde{\tau}'_k(r)}{\widetilde{\tau}_k} \right) \right|^2 \tag{2.15}$$

   where $\widetilde{\tau}'_k(r)$ is the expected $\widetilde{\tau}_{\mathrm{gw}}$ for pulsar k *assuming the source at distance r and pulsar k at the approximate distance found in step* 2.

The quantity $\psi^2(r)$ is a measure of the misfit between the observations $\widetilde{\tau}_k$ and the prediction $\widetilde{\tau}'_k$ assuming pulsars at the approximate distances and source at distance r. The r that minimizes $\psi^2(r)$ is thus an estimator for the distance to the source. Figure 2.2 shows a set of three scatter plots of the estimated r, declination θ, and right ascension ϕ relative to their actual values over 10^4 realizations of pulsar distance errors, for this example. Table 2.2 provides quantitative descriptive statistics for the distribution of errors. Clearly, if the pulsar distances can be sufficiently accurately determined, the distance and location of a detected periodic gravitational wave source can be accurately measured.

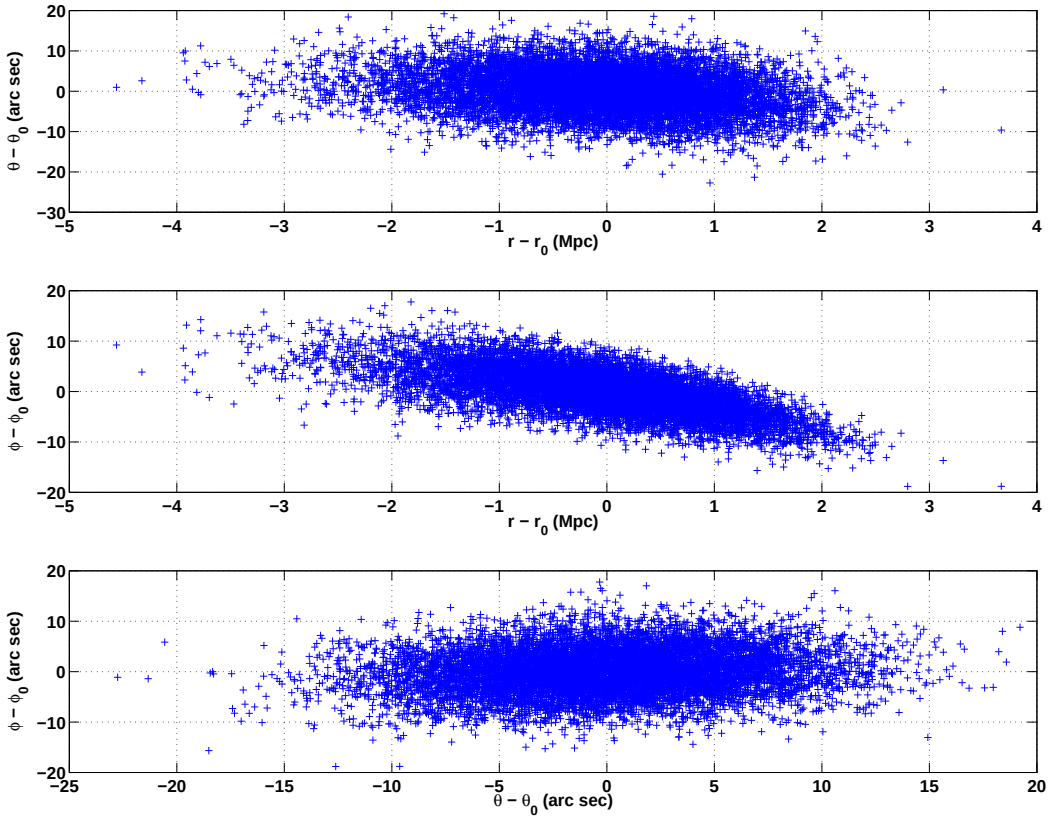


Fig. 2.2 Scatter plots of the timing parallax estimates of the distance, declination and right ascension of a simulated gravitational wave source in 3C66B (i.e., $r = 87.8$ Mpc, dec=$2^{\circ}59'31''$ and right ascension 2h23m11). The top, middle and bottom panel are respectively the distance-declination, distance-right ascension and declination-right ascension plots. See §2.3.3 for details.

Table 2.1 Four nearby millisecond pulsars monitored as part of the International Pulsar Timing Array and their current estimated distances. These four pulsars are close enough that, in the SKA era, their distances, or the distances to pulsars like them, may become known to better than 0.025 pc.

J-Name	kpc
$J0437-4715$	0.16
$J1741+1300$	0.19
$J0030+0451$	0.30
$J2124-3358$	0.25

Table 2.2 Mean, median and standard deviation of the timing parallax estimated distance and location on the sky of a simulated gravitational wave source in 3C66B. In this example the source was actually located at $r = 87.8$ Mpc, dec=$2^{\circ}59'31''$ and right ascension 2h23m11s. Only the four pulsars described in Table 2.1 were used to make these estimates. See §2.3.3 for details.

	mean	median	std. dev.
r	85.72 Mpc	85.79 Mpc	0.99 Mpc
θ	$42^{\circ}59'31''$	$42^{\circ}59'31''$	$0^{\circ}0'5.3''$
ϕ	2h23m11.0s	2h23m11.0s	0h0m0.3s

We emphasize that this is no more than a proof-of-principle demonstration. It is naive in its treatment of uncertainties in the L_k, ignores pulsar timing noise in the measured $\widetilde{\tau}_k$, presumes a sufficiently accurate knowledge of pulsar locations θ_k, and makes use of an ad hoc misfit statistic $\psi^2(r)$. Nevertheless, taken as a proof-of-principle demonstration it shows that the potential exists for pulsar timing array observations to determine the distance to periodic gravitational wave sources if the pulse arrival time dependence on the passing waves is properly modeled.

2.4 Conclusion

Gravitational waves crossing the pulsar-Earth line-of-sight lead to a disturbance in the pulsar pulse arrival time. All previous derivations of this disturbance have assumed planar gravitational wave phasefronts. The gravitational wave phasefronts from point gravitational wave sources — e.g., coalescing supermassive black hole binary systems — are curved, with curvature radius equal to the source luminosity distance. The approximation of planar wavefronts is thus valid only at distances $R \gg 2\pi f L^2/c$, where R is the Earth-source distance, L the Earth-pulsar distance, and f

the gravitational radiation frequency. For typical pulsars distances (kpc) and relevant gravitational wave frequencies ($f \lesssim \mathrm{yr}^{-1}$) the correction to the disturbance magnitude and phase owing to phasefront curvature is thus significant for sources within $\sim 100\,\mathrm{Mpc}$. Here we have derived the curved wavefront corrections to the pulsar timing response for such "nearby" sources, described when they are important, and shown that future gravitational wave observations using pulsar timing arrays may be capable of measuring luminosity distances to supermassive black hole binary systems, and other periodic gravitational wave sources, approaching or exceeding 100 Mpc.

The gravitational wave source distance measurement described here is properly considered a parallax distance measurement. The baselines over which the parallax is measured are, in this case, the timing array pulsar-Earth baselines. Crucial to the ability to observe the effects of gravitational wave phasefront curvature is a knowledge of the pulsar-Earth distance L with an uncertainty $\sigma_L \lesssim (\pi f)^{-1} \sim 2\,\mathrm{pc}\,(0.1\mathrm{yr}^{-1}/f)$. We argue that pulsar distance measurements of this accuracy, while beyond present capabilities (except in the case of J0437 − 4715, which is exceptionally bright and close), are within the capability of SKA-era observations for pulsars at distances $L \lesssim$ kpc.

When considered against the cosmic distance ladder, the distance measurement described here involves three rungs to reach distances greater than tens of Mpc: first, the determination of the astronomical unit; second, the distance to the array pulsars; finally, the distance to the gravitational wave source. Since the method described here also provides the precise source angular position the likelihood of identifying an electromagnetic counterpart (e.g., host galaxy) to the gravitational wave source is great, raising the possibility of a precise measurement of both the redshift and luminosity distance to a single object at cosmological distances, with the obvious consequences for independent verification of the parameters describing our expanding universe.

Even in the absence of an electromagnetic counterpart measurement of the luminosity distance and angular location of a supermassive black hole binary system — the most likely source

of periodic gravitational waves — enables the determination of the system's so-called "chirp mass" (Peters & Mathews, 1963) without the need to observe the binary's evolution.

We have only begun to plumb the potential of gravitational wave observations as a tool of astronomical discovery. While this potential will not be realized until gravitational waves are detected and observations become, if not routine, at least more than occasional, explorations like these will position us to more readily exploit the opportunities that future observations present.

Chapter 3

Searching for Gravitational Wave Bursts via Bayesian Nonparametric Data Analysis with Pulsar Timing Arrays

3.1 Introduction

Gravitational wave bursts are a catch-all category for signals whose durations are much shorter than the observation period. Bayesian analysis of gravitational wave detector data for gravitational wave bursts has generally modeled possible sources in terms of a finite set of parameters and their prior distributions (Seto, 2009; Pshirkov et al., 2010; Cordes & Jenet, 2012; Damour & Vilenkin, 2001; Siemens et al., 2007; Leblond et al., 2009). The posterior distribution of the parameters is then determined by the observations. From the posterior parameter distribution either a Bayes factor or signal-to-noise ratio is estimated and used to decide whether a burst has been detected (e.g. Finn & Lommen, 2010). The choice of a finite set of parameters to model the burst may impose unnecessary constraints on the form of the burst as a function of time. Additionally, the Bayes Factor may be ill-defined or overly sensitive to the parameters or their prior distribution (Gelman et al., 2004, chap. 6). Here we adopt a non-parametric approach to the problem of burst detection: instead of adopting a parameterized model for the burst and identifying parameter priors, we adopt a Gaussian process[1] prior for the burst function itself and use the observations to constrain the burst as a function of time (Rasmussen & Williams, 2006). We demonstrate our

[1] A Gaussian process is a distribution $p(f)$ over functions f, such that, for any finite subset of points in the domain of f (e.g., $\{x_1, x_2, \ldots, x_n\}$) the marginal density $p(f|\{x_1, x_2, ..., x_n\})$ is a multivariate Gaussian distribution over $\{f(x_1), f(x_2), \ldots, f(x_n)\}$ (Rasmussen & Williams, 2006, chap. 1).

analysis method by appying it to simulated pulsar timing array data including a gravitational wave burst.

Nonparametric data analysis refers to techniques that do not assume the model used to fit the data is fixed (Gibbons & Chakraborti, 2003). Bayesian nonparametrics stem from the seminal work of Ferguson (1973) and Doksum (1974), which tried to incorporate nonparametric statistics into Bayesian methodology. The advantage of Bayesian nonparametrics over other nonparametric techniques is that it is able to incorporate our expectations regarding the characteristics of the signals into a prior distribution to set a constraint on the feasible signal patterns we try to search from the data (Rasmussen & Williams, 2006). O'Hagan first proposed that the general prior distribution for regression analysis is a Gaussian process (O'Hagan, 1978). He also proved that the expected characteristics of the signal patterns can be encoded into the mean and covariance of the Gaussian process (O'Hagan, 1978). This approach has then been applied to solve various regression problems in geostatistics, meteorology, computer science, machine learning, etc (Rasmussen & Williams, 2006, chap. 2). Here we apply Bayesian nonparametric regression to search for gravitational wave bursts in the pulsar timing array data.

In Section 3.2, we describe the general principles and methodology of Bayesian nonparametric analysis. In Section 3.3 we apply this method to analyze simulated pulsar timing array data including the contribution from a gravitational wave burst. In Section 3.4 we illustrate the effectiveness of this analysis by applying it to several representative examples. Finally, we summarize our conclusions in Section 3.5.

This work is published as X. Deng, 2014, Physical Review D 90, 024020.

3.2 Bayesian nonparametric methodology

Bayesian nonparametric analysis is introduced to analyze the data when analysts cannot model the data by a set of fixed number of parameters (Walker, 2010). Here we give a brief description of the basic framework of this methodology, applied to the analysis of time series data that may include a signal we are trying to detect. We refer readers to Ghosh & Ramamoorthi (2003); Rasmussen & Williams (2006); Hjort et al. (2010) for details and other applications of Bayesian nonparametrics.

3.2.1 Framework of Bayesian nonparametric analysis

When we cannot model the signal by a fixed number of parameters, the Bayesian nonparametric approach assigns a prior distribution on an infinite dimensional space of signal functions and assigns probabilities to the functions in that space (Rasmussen & Williams, 2006, chap. 2). For example, suppose that a time series $y = \{y(t_k) : k = 1..n\}$ consists of a signal described by the function f, whose values at the t_k are $f(t_k)$, and additive zero mean noise. We can assign a prior distribution $q(\boldsymbol{f})$ on $\boldsymbol{f}$ (Finn & Lommen, 2010); then by Bayes' theorem, the posterior probability density of $\boldsymbol{f}$ would be:

$$p(\boldsymbol{f}|\mathbf{y}) \propto \Lambda(\mathbf{y}|\boldsymbol{f})q(\boldsymbol{f}) \tag{3.1}$$

where $\Lambda(\mathbf{y}|\boldsymbol{f})$ is the likelihood function, which will be a Gaussian distribution of $\mathbf{y}$ if the noise is Gaussian distributed. In our approach to Bayesian nonparametric inference, we will first choose a prior q based on our expectation of $\boldsymbol{f}$, and then infer $\boldsymbol{f}$ by computing the posterior p.

3.2.1.1 Likelihood function Λ

Our data set, which we denote $\mathbf{y} = \{\mathrm{y}_1, \mathrm{y}_2, ..., \mathrm{y}_N\}$, consists of N samples from the signal $\boldsymbol{f}$ superposed with noise, which we denote $\boldsymbol{n} = \{n_1, n_2, ..., n_N\}$. The samples are taken at times

$t = \{t_1, t_2, ..., t_N\}$. The noise is generally modeled as a Gaussian process with covariance $\mathbf{C}$, and the corresponding likelihood function is (Finn & Lommen, 2010)

$$\begin{aligned} \Lambda(\mathbf{y}|\boldsymbol{f}) &= N(\mathbf{y} - \boldsymbol{f}|\mathbf{C}) \\ &= \frac{\exp\left[-\frac{1}{2}(\mathbf{y} - \boldsymbol{f})^T \mathbf{C}^{-1}(\mathbf{y} - \boldsymbol{f})\right]}{\sqrt{(2\pi)^{\dim \mathbf{y}} \det ||\mathbf{C}||}} \end{aligned} \tag{3.2a}$$

where $\mathbf{C}$ denotes the noise covariance

$$\mathrm{C}_{lm} = \langle n_l n_m \rangle \tag{3.2b}$$

where $\langle\rangle$ denotes the ensemble average.

3.2.1.2 Prior probability density q

The probability density $q(\boldsymbol{f})$ describes our expectations of the signal before we analyze the data set $\mathbf{y}$. It plays the key role in Bayesian nonparametric analysis since it sets a constraint on the feasible function forms we try to explore (Sudderth, 2006).

In general, we may write $\boldsymbol{f}$ as a discrete Fourier transform since the observation times are discrete (Summerscales et al., 2008; Bretthorst, 1988),

$$f_\alpha = \sum_k^N \left[A_k \cos(\omega_k t_\alpha) + B_k \sin(\omega_k t_\alpha)\right] \tag{3.3}$$

for signal $\boldsymbol{f}$ beginning at time t_0 and ending at some later time $t_0 + T$, and with t_α the observation times in the interval $[t_0, t_0 + T]$. Under the minimal assumption that there is no preferred signal starting time t_0 or duration it is straightforward to find an "ignorance prior" for each of the

coefficients A_k and B_k (Bretthorst, 1988; Summerscales et al., 2008):

$$Q_k(A_k|\sigma_k) = \frac{\exp(-A_k^2/2\sigma_k^2)}{\sqrt{2\pi}\sigma_k}$$
$$Q_k(B_k|\sigma_k) = \frac{\exp(-B_k^2/2\sigma_k^2)}{\sqrt{2\pi}\sigma_k} \tag{3.4}$$

with N and σ_k undetermined.

In a conventional regression problem we would fix N and choose some prior for the σ_k (Bretthorst, 1988). Instead, however, let us take a different approach. Noting that the A_k and B_k are Gaussian random variables we may regard $\boldsymbol{f}$ as a Gaussian process with correlation function

$$\begin{aligned} \mathrm{K}_{\alpha\beta} &= \langle f_\alpha\, f_\beta \rangle \\ &= \sum_k \left[\sigma_k^2 \cos(\omega_k t_\alpha)\cos(\omega_k t_\beta) + \sigma_k^2 \sin(\omega_k t_\alpha)\sin(\omega_k t_\beta)\right] \\ &= \sum_k \sigma_k^2 \cos\left[\omega_k(t_\alpha - t_\beta)\right] \end{aligned} \tag{3.5a}$$

The covariance $\mathrm{K}_{\alpha\beta}$ is referred to as the kernel of the Gaussian process prior and it has to be positive semidefinite (Rasmussen & Williams, 2006). Correspondingly, the prior of $\boldsymbol{f}$ can be written as

$$\begin{aligned} q(\boldsymbol{f}|\theta) &= N(\boldsymbol{f}|\mathbf{K}) \\ &= \frac{\exp\left(-\frac{1}{2}\boldsymbol{f}^T\mathbf{K}^{-1}\boldsymbol{f}\right)}{\sqrt{(2\pi)^{\dim \mathbf{f}}\ \det||\mathbf{K}||}} \end{aligned} \tag{3.5b}$$

where θ denotes the unknown parameters embedded in the kernel such as σ_k. Such parameters are referred to as *hyperparameters* (Gelman et al., 2004). In full Bayesian inference, we also need to choose a prior probability density q_θ, i.e., *hyperprior*, for the hyperparameter set (Rasmussen &

Williams, 2006; Gelman et al., 2004), and the joint prior of $\boldsymbol{f}$ and the hyperparameters would be

$$q_0(\boldsymbol{f}, \theta) = q(\boldsymbol{f}|\theta)q_\theta(\theta). \tag{3.6}$$

We will infer the hyperparameters θ together with $\boldsymbol{f}$. We can incorporate our expectations regarding the possible signals $\boldsymbol{f}$ into our analysis via the prior probability density q by choosing a specific kernel and a specific hyperprior.

To recap, our non-parametric analysis has characterized the signal by a function f drawn from an infinite dimensional function space characterized by the Gaussian process prior with kernel $\mathbf{K}(\theta)$. Unlike conventional regression approaches there is no predetermined mathematical form of $\boldsymbol{f}$ aside from those we impose via a prior on the kernel $\mathbf{K}$. An appropriate choice of $\mathbf{K}$ can set a strong constraint on the smoothness, the trend and the variations of the signal patterns (Rasmussen & Williams, 2006). A similar approach has been introduced by Finn & Lommen (2010). However, the prior $q(\boldsymbol{f})$ they choose contains no prior information of the temporal correlation of the signal, which ignores the imporant characteristics of the signal patterns.

From the discussion above, we can see that characterizing the signal by the parameter set $\boldsymbol{f}$ and setting its prior distribution by a Gaussian process Eq. (3.5b) is equivalent to characterizing the signal by the superposition of a flexible number of basis functions such as discrete Fourier transforms and setting Gaussian prior distributions for the corresponding coefficients of those bases.

3.2.1.3 Bayesian nonparametric inference

Since we have chosen the appropriate priors, we can make the inference of $\boldsymbol{f}$, i.e., the function form of the signal. Assuming the noises are Gaussian distributed with zero mean, the joint posterior

probability density of $\boldsymbol{f}$ and hyperparameters would be

$$\begin{aligned} p(\boldsymbol{f},\theta|\mathbf{y}) &= \frac{1}{Z(\mathbf{y})}\Lambda(\mathbf{y}|\boldsymbol{f})q(\boldsymbol{f}|\theta)q_\theta(\theta) \\ &= \sqrt{\frac{\det||\mathbf{A}||}{(2\pi)^{\dim\mathbf{A}}}}\exp\left[-\frac{1}{2}(\boldsymbol{f}-\boldsymbol{f}_m)^T\mathbf{A}(\boldsymbol{f}-\boldsymbol{f}_m)\right] \\ &\times \frac{1}{Z(\mathbf{y})}\Lambda_\theta(\mathbf{y}|\theta)q_\theta(\theta) \end{aligned} \tag{3.7a}$$

where $Z(\mathbf{y})$ is the normalization constant; $\mathbf{A}$ is

$$\mathbf{A} = \mathbf{K}^{-1} + \mathbf{C}^{-1} \tag{3.7b}$$

with $\mathbf{C}$ denoting the noise covariance matrix; and $\boldsymbol{f}_m$ satisfies

$$\mathbf{A}\boldsymbol{f}_m = \mathbf{C}^{-1}\mathbf{y}. \tag{3.7c}$$

$\Lambda_\theta(\mathbf{y}|\theta)$ is the likelihood function of hyperparameters after marginalizing over $\boldsymbol{f}$

$$\begin{aligned} \Lambda_\theta(\mathbf{y}|\theta) &= \int \Lambda(\mathbf{y}|\boldsymbol{f})q(\boldsymbol{f}|\theta)\,\mathrm{d}\boldsymbol{f} \\ &= \frac{\exp\left[-\frac{1}{2}\mathbf{y}^T\mathbf{C}^{-1}\mathbf{y}\right]}{\sqrt{(2\pi)^{\dim\mathbf{y}}\det||\mathbf{C}||}} \times \frac{\exp\left[\frac{1}{2}(\mathbf{C}^{-1}\mathbf{y})^T\mathbf{A}^{-1}(\mathbf{C}^{-1}\mathbf{y})\right]}{\sqrt{\det||\mathbf{A}||\det||\mathbf{K}||}}. \end{aligned} \tag{3.7d}$$

If we would like to infer $\boldsymbol{f}$, we need to choose a hyperprior q_θ and marginalize over hyperparameter θ to obtain the marginalized posterior for $\boldsymbol{f}$, i.e.,

$$p_\theta(\boldsymbol{f}|\mathbf{y}) = \int p(\boldsymbol{f},\theta|\mathbf{y})\,\mathrm{d}\theta \tag{3.7e}$$

and if we are also interested in estimating hyperparameter θ, we need to marginalize over $\boldsymbol{f}$ and obtain the posterior for θ, i.e.,

$$p_\theta(\theta|\mathbf{y}) = \int p(\boldsymbol{f}, \theta|\mathbf{y})\,\mathrm{d}\boldsymbol{f} = \frac{1}{Z(\mathbf{y})}\Lambda_\theta(\mathbf{y}|\theta)q_\theta(\theta). \tag{3.7f}$$

Eq. (3.7) summarizes Bayesian nonparametric inference, which gives estimates of the signal $\boldsymbol{f}$ and hyperparameters.

3.2.2 Comparison with Bayesian parametric analysis

At this point, it is worth comparing the Bayesian nonparametric inference described above and the conventional inference methods that try to fit the data by a model with a fixed number of parameters.

The conventional inference methods assume that we know the analytical formulae of the signals we try to detect and we can characterize them with a fixed number of parameters. For example, to detect gravitational waves from a non-spinning binary black hole inspiral, we can characterize the signal by a post-Newtonian formula with 8 parameters to characterize the signal — chirp mass, mass ratio, coalescence time, strain amplitude, sky location of the source, polarization and inclination angle of the orbital plane (Blanchet, 2006). In this way, we can obtain key characteristics of the signals since the parameters we use to model the signal usually have explicit physical meanings. However, this approach requires that we are able to determine the analytical formulae of the signals.

When the analytical formulae of the signals are not available, the conventional approach could model the signal by a linear superposition of a finite number of basis functions (Gelman

et al., 2004, chap. 16), i.e.,

$$f_\alpha = \sum_k^N \alpha_k \Phi_k(t_\alpha) \tag{3.8}$$

where Φ_k is the basis function and α_k is the corresponding coefficient that would be the unknown parameters we try to infer; N is the number of the basis functions, which has to be fixed. For example, we can choose Φ_k as the Fourier modes like Eq. (3.3). However, if we do not know much information about the signals, we do not know how many basis functions we need to choose. A model with too many basis functions would be so complex as to overfit the data while a model with too few basis functions would be so simple as to underfit the data.

In contrast, Baysian nonparametric approach directly infers the signal pattern by assigning Gaussian process prior distributions to set strong constraints on it, which would avoid using a detailed physical model. However, since Bayesian nonparametrics do not use the parameters that describe the physical characteristics of the signal like conventional methods, it would provide less direct insight for the signal.

3.3 Bayesian nonparametric analysis on gravitational wave bursts with pulsar timing arrays

To seek evidence of a gravitational wave burst, we need to match the timing residuals with a model that characterizes the contribution of the burst. However in general cases, we do not have physical models for the burst sources and so we are not able to characterize the gravitational wave burst by some analytical formula or by a fixed number of basis functions. Therefore, we introduce the Bayesian nonparametric analysis described in the previous section to detect and characterize gravitational wave bursts with a pulsar timing array.

3.3.1 Properties of the pulsar timing response to the passage of a gravitational wave burst

When the duration of the gravitational wave bursts ΔT and the observational duration of pulsar timing array data T are much less than $L(1+\hat{k}\cdot\hat{n})$, it is most likely that only the Earth term contributes to the correlated timing residuals (Finn & Lommen, 2010), unless (1) there are fortuitous lines of sight of a pair or more of the pulsars where the time delay among those pulsar terms is small enough to be within the observational duration (Pitkin, 2012); (2) there are a large number of pulsars providing a long time baseline that might be able to detect more burst sources (Cordes & Jenet, 2012). However, in the current IPTA, there are not so many pulsars whose timing noises are low enough to be effective in gravitational wave detection (Hobbs et al., 2010; Demorest et al., 2012; Manchester et al., 2013), and it is uncertain when we can have a pulsar timing array with a large number of low timing noise pulsars. Therefore, in this chapter, we only consider the effects of the Earth term on the detection of gravitational wave bursts, and the corresponding pulsar timing response can be written as

$$\tau_j(t) = F_j^{(+)}\tau_{(+)}(t) + F_j^{(\times)}\tau_{(\times)}(t) \tag{3.9a}$$

where $F_j^{(A)}$ is the pattern function of the jth pulsar,

$$F_j^{(A)} = -\frac{\hat{n}_j^l \hat{n}_j^m e_{lm}^{(A)}}{2(1+\hat{k}\cdot\hat{n}_j)}. \tag{3.9b}$$

In the pulsar timing array waveband, the internal motion of the gravitational wave sources is expected to be smooth with the evolution of time (Hobbs et al., 2010). Correspondingly, the

waveform $h_{(A)}(t)$ of a gravitational wave burst is expected to be a smooth funtion of the observation times (Misner et al., 1973) and thus $\tau_{(A)}(t)$ should also be a smooth function of time.

Furthermore, for a detectable gravitational wave burst through pulsar timing arrays, the duration of the burst should be shorter than the observation duration, or otherwise it is not considered a "burst" signal.

These two properties of the timing residual are very important and they are considered as our prior knowledge of the gravitational wave bursts. As we will see in the next subsection, we will take advantage of this prior information to model the pulsar timing response to the passage of a gravitational wave burst.

3.3.2 The choice of prior probability distribution

3.3.2.1 Priors of $\tau_{(+)}$ and $\tau_{(\times)}$

As described in Section 3.2, to apply Bayesian nonparametrics to detect gravitational wave bursts, we need to choose an appropriate prior probability to constrain feasible function forms of the signals. For $\tau_{(+)}$ and $\tau_{(\times)}$, we have 3 expectations as discussed in the last subsection:

(1) They are the same for different pulsars.

(2) They are smooth on the timescales long compared to the inter-sample time.

(3) They have a characteristic time duration that is shorter than the observation time.

We need to choose their Gaussian process priors with appropriate kernels to fulfill these three expectations.

We do not have any information on the arrival time of the gravitational wave burst, which means that priors of $\tau_{(+)}$ and $\tau_{(\times)}$ should hold the time translational symmetry (Summerscales et al., 2008) and the kernels should only depend on the difference between observation times,

i.e., stationary (Rasmussen & Williams, 2006). $\tau_{(+)}$ and $\tau_{(\times)}$ are also expected to be infinitely differentiable on time; correspondingly, their mean squares under their Gaussian process priors have to be infinitely differentiable, which requires their stationary kernels $\mathrm{K}_{(A)}(\Delta t)$ to be infinitely differentiable at $\Delta t = 0$ (Adler, 1981, chap. 2), where Δt denotes the difference between any two observation times. This is a very rigorous requirement since the stationary kernels have to be both positive semidefinite and infinitely differentiable, and only a few kernels we know satisfy this (Rasmussen & Williams, 2006, chap. 4). The one with the least number of hyperparameters is the square exponential kernel (Rasmussen & Williams, 2006, chap. 4)

$$\mathrm{K}_{(+)}(\Delta t) = \sigma_{+}^{2} \exp(-\frac{\Delta t^{2}}{2\,\lambda^{2}}) \tag{3.10}$$

$$\mathrm{K}_{(\times)}(\Delta t) = \sigma_{\times}^{2} \exp(-\frac{\Delta t^{2}}{2\,\lambda^{2}}), \tag{3.11}$$

where σ_{+}, $\sigma_{\times}$ and λ are hyperparameters. We assign two different kernels to two polarization components of the gravitational wave burst because the two components are independent of each other (Finn & Lommen, 2010). The hyperparameters σ_{+} and $\sigma_{\times}$ characterize the rms amplitudes of the two polarization components. The hyperparameter λ is the characteristic time scale of the burst, which characterizes temporal correlation of $\tau_{(+)}$ and $\tau_{(\times)}$ along the observation times. When $\Delta t > \sqrt{2}\lambda$, the values of the kernels would exponentially damp, and $\tau_{(A)}(t)$ and $\tau_{(A)}(t + \Delta t)$ would be almost uncorrelated. Correspondingly, $\sqrt{2}\lambda$ should be approximately the duration of the burst and λ should be shorter than the observation duration. Here we assume that λ is the same for both of the two polarization components because the duration of the burst scales with the ratio of the typical radius to the internal velocity of the source (Finn & Lommen, 2010), which is the same for the two components.

To recap, the priors of $\tau_{(+)}$ and $\tau_{(\times)}$ would be two zero mean Gaussian processes with kernels in Eq. (3.10). They fulfill the three expectations stated above as:

(1) The priors of $\tau_{(+,\times)}$ are independent of pulsars.

(2) Kernels in Eq. (3.10) guarantee that $\tau_{(+,\times)}$ sampled from the Gaussian process priors are most likely smooth functions of time since the kernels are infinitely differentiable at $\Delta t = 0$.

(3) The hyperparameter λ in kernels of Eq. (3.10) characterizes the characteristic time-scale of the burst.

A similar approach has been introduced by Finn & Lommen (2010). However, the prior $q(\boldsymbol{f})$ they choose contains no prior information of the temporal correlation of the signal, i.e., no parameter characterizing the burst duration. Compared with Finn & Lommen (2010), our method can instead infer the burst duration, one of the most important physical parameters of the burst, which shows the key power of the Bayesian nonparametric methods.

3.3.2.2 Priors of timing residuals induced by a gravitational wave burst

Now we need to determine a prior probability density for the pulsar timing residuals induced by a gravitational wave burst. We denote $\boldsymbol{\tau}$ as the timing residuals for all the pulsars in a pulsar timing array, i.e., $\boldsymbol{\tau} = \{\tau_1, \tau_2, ..., \tau_j, ...\}$. Because $\boldsymbol{\tau}$ is the linear superposition of $\tau_{(+)}$ and $\tau_{(\times)}$ (see Eq. (3.9)), whose priors are independent zero mean Gaussian process priors, the prior of $\boldsymbol{\tau}$ should also be a zero mean Gaussian process prior, and its kernel should be the linear superposition of the two kernels of the Gaussian process priors of $\tau_{(+)}$ and $\tau_{(\times)}$, i.e.,

$$q(\boldsymbol{\tau}|\hat{k}) = \frac{\exp\left[-\frac{1}{2}\boldsymbol{\tau}^T\mathbf{K}^{-1}\boldsymbol{\tau}\right]}{\sqrt{(2\pi)^{\dim\mathbf{K}}\det||\mathbf{K}||}} \tag{3.12a}$$

where $\mathbf{K}$ is expressed as

$$\mathrm{K}_{j(\alpha),\,k(\beta)} = \left(\sigma_{+}^{2} F_{j}^{(+)} F_{k}^{(+)} + \sigma_{\times}^{2} F_{j}^{(\times)} F_{k}^{(\times)}\right) \exp\left[-\frac{\left(t_{j(\alpha)} - t_{k(\beta)}\right)^{2}}{2\,\lambda^{2}}\right] \tag{3.12b}$$

where j, k are pulsar indices and α, β are the indices for the observation times of pulsar timing measurements. Here we have used Eq. (3.9) and Eq. (3.10).

3.3.2.3 Prior of hyperparameters

The prior probability density of $\boldsymbol{\tau}$, i.e., Eq. (3.12), contains several hyperparameters. We also need to choose appropriate prior probability densities for those hyperparameters. According to Eq. (3.12), we have 5 hyperparameters — rms gravitational wave amplitudes σ_{+} and $\sigma_{\times}$, characterisitic time-scale λ, and gravitational wave propagation direction $\hat{k}$ which is embedded in the pattern function $F^{(+)}$ and $F^{(\times)}$.

We first choose the hyperpriors for σ_{+} and $\sigma_{\times}$. The pulsar timing array data is assumed to be normally distributed with the mean $\boldsymbol{\tau}$ due to the assumption that the noises are normally distributed, and the prior distribution of $\boldsymbol{\tau}$ is also chosen to be a Gaussian process. This is a two-level normal model and the choice of hyperprior on the rms amplitude is discussed in detail in (Gelman, 2006). Both σ_{+} and $\sigma_{\times}$ appear like scale parameters, suggesting the use of the Jeffreys prior (Jeffreys, 1946); however, the Jeffreys prior leads to a non-normalizable posterior probability density (Gelman, 2006), which means it is not an appropriate choice for hyperpriors of σ_{+} and $\sigma_{\times}$. If the dimension of $\boldsymbol{\tau}$, which in our case equals the number of data points, is larger than 5, a uniform hyperprior distribution of the rms amplitudes is recommended (Gelman, 2006), i.e.,

$$q_{+}(\sigma_{+}), q_{\times}(\sigma_{\times}) \propto 1. \tag{3.13a}$$

We then choose the hyperprior for the characteristic time-scale λ. We expect λ should be shorter than the observation duration T. As discussed before, $\sqrt{2}\lambda$ characterizes the duration of the burst, so if λ is much longer than T, the gravitational wave signal cannot be considered as a "burst". To make sure the posterior distribution is normalizable, we assume the hyperprior of λ is a proper uniform distribution from 0 to T,

$$q_{\lambda}(\lambda) = \frac{1}{T}. \tag{3.13b}$$

We finally set the hyperprior for the sky location of the source. We expect the sources of gravitational wave bursts to be uniformly distributed across the sky, so the prior of the gravitational wave propagation direction $\hat{k}$ is

$$q_{\hat{k}}(\hat{k}) = \frac{1}{4\pi}. \tag{3.13c}$$

3.3.3 Inferring $\boldsymbol{\tau}$ and hyperparameters

Having chosen the prior probability distribution of $\boldsymbol{\tau}$, we need to write down its probability posterior density to make a Bayesian nonparametric inference, as described in Section 3.2.1. We denote the observed timing residuals of the jth pulsar as d_j and the contribution by gravitational wave burst as τ_j, then the likelihood function of the jth pulsar is

$$\begin{aligned}\Lambda(\mathrm{d}_j|\tau_j) &= N(\mathrm{d}_j - \tau_j|\mathrm{C}_j) \\ &= \frac{\exp\left[-\frac{1}{2}(\mathrm{d}_j - \tau_j)^T \mathrm{C}_j^{-1}(\mathrm{d}_j - \tau_j)\right]}{\sqrt{(2\pi)^{\dim \mathrm{d}_j} \det ||\mathrm{C}_j||}}\end{aligned} \tag{3.14a}$$

where C_j is the noise covariance of the jth pulsar. We also assume that the timing noises are uncorrelated among different pulsars. Correspondingly, for a pulsar timing array composed of N_p

pulsars, the likelihood function of its timing residuals $\mathbf{d}$ is

$$\Lambda(\mathbf{d}|\boldsymbol{\tau}) = \prod_{j=1}^{N_p} \Lambda(\mathrm{d}_j|\tau_j)$$

$$= N(\mathbf{d}-\boldsymbol{\tau}|\mathbf{C}). \tag{3.14b}$$

Following the discussion in Section 3.2.1.3 with Eq. (3.14b), Eq. (3.12) and Eq. (3.13), we can determine the joint posterior probability density of $\boldsymbol{\tau}$ and hyperparameters,

$$\begin{aligned} p(\boldsymbol{\tau},\hat{k},\sigma_+,\sigma_\times,\lambda|\mathbf{d}) &= \frac{1}{Z(\mathbf{d})}\Lambda(\mathbf{d}|\boldsymbol{\tau})q(\boldsymbol{\tau}|\hat{k},\sigma_+,\sigma_\times,\lambda)q_{\hat{k}}(\hat{k})q_+(\sigma_+)q_\times(\sigma_\times)q_\lambda(\lambda) \\ &= \sqrt{\frac{\det||\mathbf{A}||}{(2\pi)^{\dim\mathbf{A}}}}\exp\left[-\frac{1}{2}(\boldsymbol{\tau}-\boldsymbol{\tau}_m)^T\mathbf{A}(\boldsymbol{\tau}-\boldsymbol{\tau}_m)\right] \\ &\times \frac{1}{Z(\mathbf{d})}\Lambda_\theta(\mathbf{d}|\hat{k},\sigma_+,\sigma_\times,\lambda)q_{\hat{k}}(\hat{k})q_+(\sigma_+)q_\times(\sigma_\times)q_\lambda(\lambda) \end{aligned} \tag{3.15a}$$

where $Z(\mathbf{d})$ is the normalization constant and $\mathbf{A}$ is

$$\mathbf{A} = \mathbf{K}^{-1} + \mathbf{C}^{-1} \tag{3.15b}$$

and $\boldsymbol{\tau}_m$ satisfies

$$\mathbf{A}\boldsymbol{\tau}_m = \mathbf{C}^{-1}\mathbf{d}. \tag{3.15c}$$

$\Lambda_\theta(\mathbf{d}|\hat{k},\sigma_+,\sigma_\times,\lambda)$ is given by

$$\begin{aligned} \Lambda_\theta(\mathbf{d}|\hat{k},\sigma_+,\sigma_\times,\lambda) &= \int \Lambda(\mathbf{d}|\boldsymbol{\tau})q(\boldsymbol{\tau}|\hat{k},\sigma_+,\sigma_\times,\lambda)\,\mathrm{d}\boldsymbol{\tau} \\ &= \frac{\exp\left[-\frac{1}{2}\mathbf{d}^T\mathbf{C}^{-1}\mathbf{d}\right]}{\sqrt{(2\pi)^{\dim\mathbf{d}}\det||\mathbf{C}||}} \times \frac{\exp\left[\frac{1}{2}(\mathbf{C}^{-1}\mathbf{d})^T\mathbf{A}^{-1}(\mathbf{C}^{-1}\mathbf{d})\right]}{\sqrt{\det||\mathbf{A}||\det||\mathbf{K}||}} \end{aligned} \tag{3.15d}$$

Eq. (3.15) summarizes the Bayesian nonparametric inference, which gives an estimate of the time series function $\boldsymbol{\tau}$ and hyperparameters. Compared with Eq. (23) in Finn & Lommen (2010), our kernel $\mathbf{K}$ is a non-diagonal matrix rather than diagonal, which provides an important constraint on the temporal correlation of the signal.

3.4 Examples

3.4.1 Overview

To illustrate the effectiveness of the analysis techniques just described, we apply them to simulated observations of a gravitational wave burst generated by a periapsis passage of a long period supermassive black hole binary in a highly eccentric orbit. We consider 3 cases:

(1) A strong signal, in which we can not only detect the signal, but also localize the source in the sky.

(2) A weak signal, in which we are able to detect the signal, but not able to accurately localize its source.

(3) No signal at all.

For these examples, we use 4 pulsars in the current IPTA (Hobbs et al., 2010; Demorest et al., 2012; Manchester et al., 2013) which are the most accurately timed as described in Table 3.1. The capability of detecting and characterizing gravitational waves is dominated by these best pulsars, although they are the minority of the full IPTA set (Burt et al., 2011). I use this pulsar timing array for all the simulated data sets in this thesis. The timing noises of those 4 pulsars are a superposition of short-timescale white noise with rms timing residual given in Table 3.1 and long-timescale red noise normalized to have the same spectral density as the white noise at a frequency of 0.2 yr^{-1}.

The data sets we use for these examples are constructed by:

Table 3.1 Four IPTA pulsars we use, their white timing noise rms and the telescopes from which the timing residuals are measured (Hobbs et al., 2010; Demorest et al., 2012; Manchester et al., 2013)

Pulsar	rms Residual (ns)	Telescope
J1713+0747	30	AO
J1909−3744	38	GBT
J0437−4715	75	Parkes
J1857+0943	111	AO

(1) Evaluating the pulsar timing response of each pulsar to a passage of a gravitational wave generated by a periapsis passage of a long period supermassive black hole binary in highly eccentric orbit (see Section 3.4.2.1).

(2) Adding the pulsar timing noise to the gravitational wave signals obtained by the first step (see Section 3.4.2.2).

(3) Removing the best-fit linear trend from the noisy timing residuals obtained by the second step because in reality, the linear trend in the data cannot be distinguished from the systematic effects caused by pulsar spin and spin down (Hobbs et al., 2006; Edwards et al., 2006).

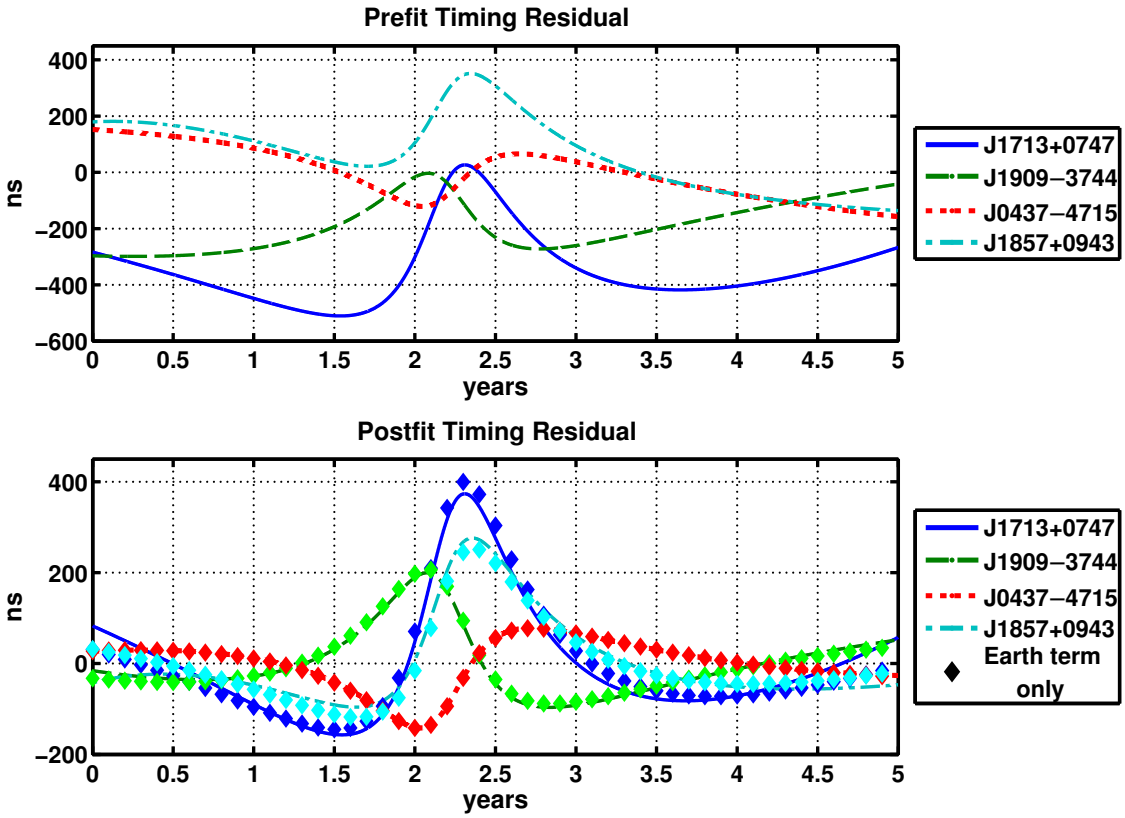
3.4.2 Construction of simulated data sets

3.4.2.1 Periapsis passage of a long period supermassive black hole binary in a highly eccentric orbit

Our hypothetical gravitational wave burst source is the periapsis passage of a supermassive black hole binary with a total mass of $2 \times 10^9\,\mathrm{M}_\odot$, a symmetric mass ratio of 0.2, a period of 20 years, an eccentricity of 0.8, and an orbital inclination angle of 30°. The duration of this burst is 0.42 yr, estimated as twice the ratio of the impact parameter to the velocity at periapsis (Finn & Lommen, 2010). It is in the direction of the Virgo cluster (RA 12h27m, dec $12^\circ 43'$) and we change the luminosity distance to obtain the strong and weak signals. We randomly sample 50 observation

times that are uniformly distributed across the 5 year observations for each pulsar. Fig. 3.1 shows the gravitational wave induced timing residuals of 4 pulsars when the source is at a distance of 16.5 Mpc (in the Virgo cluster).

Fig. 3.1 Prefit and postfit pulsar timing residuals induced by a gravitational wave burst



Notes. In this example, the gravitational wave burst is generated by the periapsis passage of a supermassive black hole binary with a total mass of $2 \times 10^9\,\mathrm{M}_\odot$, a symmetric mass ratio of 0.2, a period of 20 years, an eccentricity of 0.8, an orbital inclination angle of 30°, and located in the Virgo Cluster. The upper panel shows the timing residuals of the 4 IPTA pulsars induced by this burst, and the lower panel shows the same timing residuals but after "fitting-out" the linear trend. The "diamond" plots in the lower panel show the timing residuals only considering Earth term contributions.

3.4.2.2 Pulsar timing noise

The millisecond pulsars used in the current international pulsar timing array typically show white noise on short timescales, and few of them turn to red noise on timescales $\gtrsim 5$ years (Demorest et al., 2012; Manchester et al., 2013). For demonstrations here, we model the timing noise as the superposition of white noise and red noise, with the red noise contribution normalized to have the same amplitude as the white noise contribution at the frequency $f_r = 0.2\,\mathrm{yr}^{-1}$. The power spectral density is taken to be

$$S_n(f) = \sigma_n^2 + \sigma_n^2 \left[\frac{1 + \left(\frac{f}{f_0}\right)^2}{1 + \left(\frac{f_r}{f_0}\right)^2}\right]^{-5/2} \tag{3.16a}$$

where

$$\sigma_n = (\text{white noise rms}) \tag{3.16b}$$

$$f_r = \left(\text{red-white noise crossover frequency, } 0.2\,\mathrm{yr}^{-1}\right) \tag{3.16c}$$

and f_0 softens the noise spectrum at ultra-low frequency. As long as f_0 is much less than the pulsar timing array frequency band, its value does not matter. In the simulation we set f_0 equal to 0.01nHz. We choose the power index of the red noise spectrum as -5 because the few millisecond pulsars showing red noises have a noise spectrum with this feature (Shannon & Cordes, 2010).

The covariance matrix of the noise will be the Fourier transform of the noise spectrum density Eq. (3.16a) to time domain, i.e.,

$$\mathrm{C}(t_i, t_j) = \sigma_n^2 \left(\delta_{ij} + \frac{8\pi^2}{3}\left[1 + \left(\frac{f_r}{f_0}\right)^2\right]^{5/2} f_0^3 (t_i - t_j)^2 K_2(2\pi f_0 |t_i - t_j|)\right) \tag{3.17}$$

where $t_{i,j}$ are the "observation times" and K_2 is the modified Bessel function of the second kind with index 2. The pulsar timing noises for each pulsar are sampled from multivariate normal

distribution with zero mean and covariance matrix Eq. (3.17). For the rest of this thesis, I use this kind of pulsar timing noise for all the simulated pulsar timing data sets.

3.4.3 Analysis of simulated data sets

Our Bayesian nonparametric analysis is designed to investigate whether a gravitational wave burst is present in the simulated dataset, and also to infer the source sky location, the burst duration and other hyperparameters. We have found that for weak signals, it may be able to clearly detect the gravitational wave but not to accurately localize the sources or infer $\boldsymbol{\tau}$; but for strong signals, it may be possible to both detect the gravitational wave and also precisely localize the sources and infer $\boldsymbol{\tau}$. We illustrate this point in the following three subsections, first investigating the strong signal example, then weak signal counterpart, and finally, we analyze the dataset consisting of timing noises alone for comparative study. The results are summarized in Table 3.2. The signal-to-noise ratio (SNR) is defined as that in Finn & Lommen (2010), i.e.,

$$\mathrm{SNR}^2 = \boldsymbol{\tau}_0 \mathbf{C}^{-1} \boldsymbol{\tau}_\mathbf{0} \tag{3.18}$$

where $\boldsymbol{\tau}_0$ denotes the "actual" signal and $\mathbf{C}^{-1}$ is the noise covariance matrix.

3.4.3.1 Strong signal

We simulate the strong signal by placing the source described in Section 3.4.2.1 at a distance of 16.5Mpc (in the Virgo Cluster). The peak gravitational wave induced timing residuals of the "+" and "×" polarization components are respectively 133 ns and 131 ns. The duration of this burst, which may be estimated as twice the ratio of the impact parameter to the velocity at the periapsis of the elliptical orbit (Finn & Lommen, 2010), is about 0.42 yr. We apply our Bayesian

Table 3.2 Results for the Bayesian nonparametric analysis on 3 simulated data sets

SNR	ΔDIC	$\Delta\Omega_k$ (deg^2)	ϵ_{σ_+}	$\epsilon_{\sigma_\times}$	ϵ_λ	max Z-score
30	-956	3600	23.6%	32.4%	11.4%	0.290
5	-14	27000	88.1%	112.4%	50.0%	0.294
0	5	40800	174.0%	232.0%	58.9%	0.332

Notes. In all cases, the signal corresponds to a gravitational wave burst generated from the periapsis passage of a supermassive black hole binary with a total mass of $2 \times 10^9\,M_\odot$, a symmetric mass ratio of 0.2, period of 20 years, an eccentricity of 0.8, and an orbital inclination angle of 30°, propagating from the direction of the Virgo cluster. The source is placed at 16.5Mpc for simulating a strong signal and at 100Mpc for simulating a weak signal. $\Delta\Omega_k$ denotes the measured uncertainty of the sky location of the source, which takes 99% of the total probability; ϵ_{σ_+}, $\epsilon_{\sigma_\times}$ and ϵ_λ respectively denote the fractional uncertainty of σ_+, $\sigma_\times$ and λ (fractional uncertainty being defined as the measured uncertainty that takes 68.5% of the total probability over the mean value of the measured parameter). The last column "max Z-score" denotes the maximum absolute value of the Z-score of the MCMC samples of the three hyperparameters for each case, which implies that our MCMC samples are all reasonably convergent.

nonparametric analysis described in Section 3.2 to this "strong signal" data set (see Fig. 3.2), and the first row of Table 3.2 and Fig. 3.3 and 3.4 summarize the results:

(1) From the first row of Table 3.2, we see that the difference between the DICs of the positive hypothesis and null hypothesis is -956, corresponding to decisive evidence for the presence of a gravitational wave burst in the data set.

(2) Having concluded that a signal is present, we use the Bayesian nonparametric inference desccribed in Section 3.3.3 to infer the sky location of the gravitational wave source. Fig. 3.3 shows the posterior probability density of the sky location $\hat{k}$ marginalizing over all possible $\boldsymbol{\tau}$ and all other hyperparameters. The area corresponding to 99% of the total probability is $3600\,\mathrm{deg}^2$, and the actual sky location of the source is within this area.

(3) We can also infer other hyperparameters such as rms gravitational wave amplitude $\sigma_{+,\times}$ and characteristic length-scale λ. Fig. 3.4 shows the posterior probability density of these hyperparameters, marginalized over all possible $\boldsymbol{\tau}$ and sky locations. The mean value of σ_+

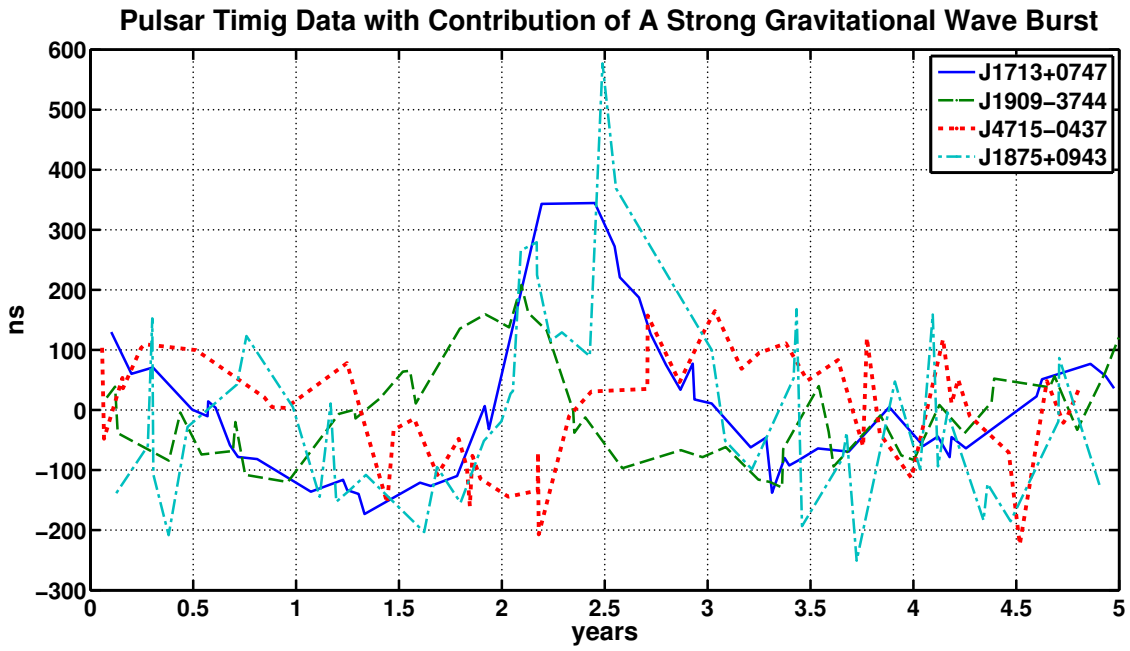


Fig. 3.2 Simulated pulsar timing data set containing a strong burst.

is 114.5 ns and that of $\sigma_\times$ is 119.8 ns. Their rms errors are respectively 27.1 ns and 38.8 ns. Correspondingly, their fractional uncertainties are respectively 23.6% and 32.4%, and the "actual" peak gravitational wave induced timing residuals of "+" polarization component (133 ns) and "×" polarization component (131 ns) are both within the uncertainty interval. For λ, the mean value and rms error are respectively 0.32 yr and 0.037 yr. Correspondingly the fractional uncertainty is 11.4% and the actual duration of the burst divided by $\sqrt{2}$ (0.297 yr), which is the characteristic length-scale (see discussion in Section 3.3.2.1), also lies within the uncertainty interval.

Therefore, for this strong signal, we can find decisive evidence of its presence in the simulated data, and also we can localize its source and estimate its burst duration well.

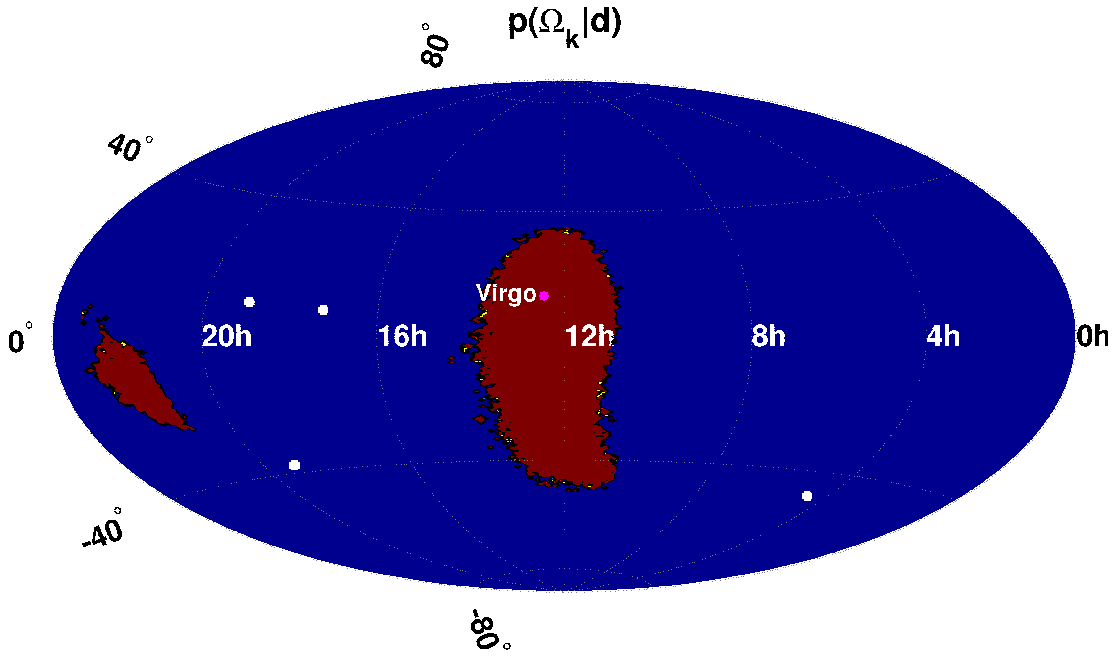


Fig. 3.3 Posterior probability density that the source is found at location Ω_k for the analysis of the "strong signal" data set. The inferred sky location has a 99% of probability staying within the red region, and the actual sky location is labelled as "Virgo". The white squares show the locations of the 4 IPTA pulsars used as our pulsar timing array. See main text for details.

3.4.3.2 Weak signal

To simulate a weak signal, we place the gravitational wave source at a distance of 100 Mpc but still keep it in the direction of the Virgo Cluster. The peak gravitational wave induced timing residuals of the two polarization components are respectively 21.8 ns and 21.6 ns. We apply the Bayesian nonparametric analysis to this "weak signal" data set (see Fig. 3.5), and the second row of Table 3.2 and Fig. 3.6 and 3.7 summarize the results:

(1) From the second row of Table. 3.2, we see that the difference between the DICs of the two exclusive hypotheses is -14, which implies strong evidence of the presence of a gravitational wave burst in the data set.

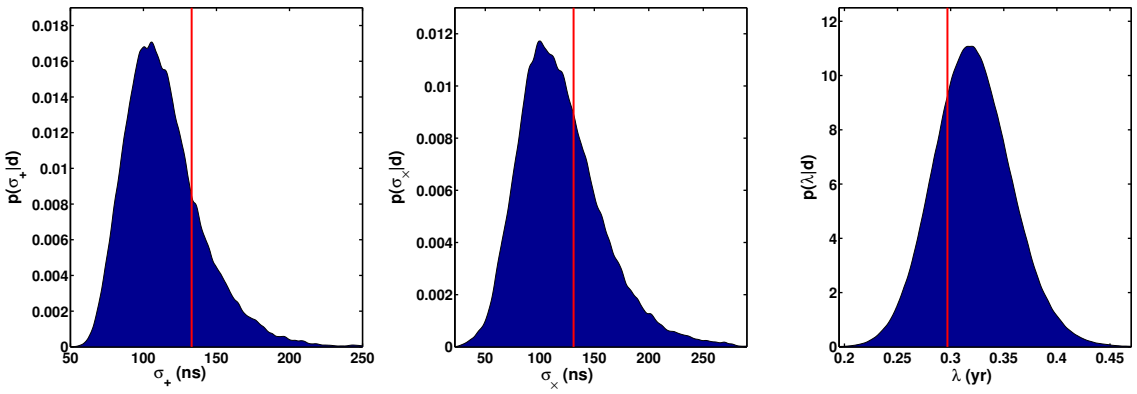


Fig. 3.4 Posterior probability densities of 3 hyperparameters — σ_{+}, $\sigma_{\times}$ and λ, for analysis for the "strong signal" data described in Section 3.4.3.1. The left panel shows that the fractional uncertainty of σ_{+}, i.e., its standard deviation over its mean value, is about 23.7%, and the red line represents the peak gravitational wave induced timing residual of the "+" polarization component of the simulated source. The middle panel shows that the fractional uncertainty of $\sigma_{\times}$ is about 32.4%, and the red line represents the peak gravitational wave induced timing residual of the "×" polarization component of the simulated source. The right panel shows that the fractional uncertainty of λ is about 11.4%, and the red line represents the duration of the simulated burst divided by $\sqrt{2}$.

(2) Having obtained strong evidence for the presence of a signal, we begin to infer the sky location of the source. Fig. 3.6 shows the posterior probability density of the sky location. The area that contains 99% of the total probability is about $27000\,\mathrm{deg}^2$.

(3) We also estimate other hyperparameters. Fig. 3.7 shows the posterior probability densities of $\sigma_{+,\times}$ and λ.

Therefore, for this weak signal, we can only marginally detect it and make crude inferences.

3.4.3.3 No signal

For comparative study, we also apply our Bayesian nonparametrics analysis to a data set with timing noises only. The third row of Table 3.2 and Fig. 3.8 summarize the results. The difference between the DICs of the two exclusive hypotheses is 5, which shows that the data favors the null hypothesis.

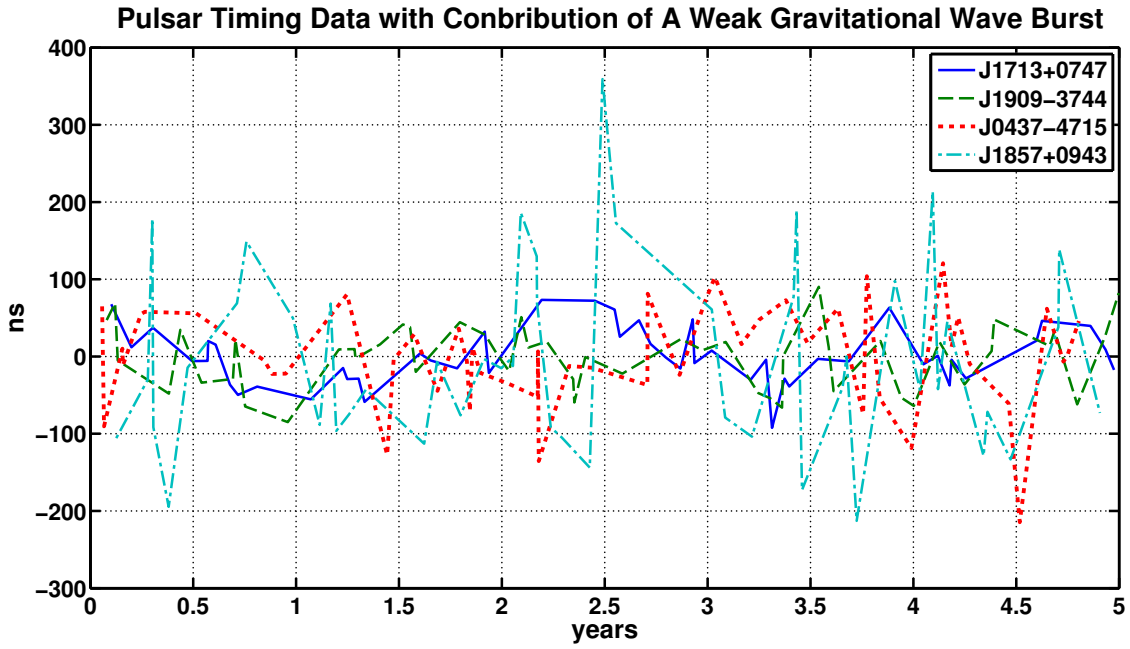


Fig. 3.5 Simulated data set containing a weak burst signal.

To compare the DIC analysis with the false alarm analysis in frequentist analysis, I repetitively compute the DIC differences (ΔDICs) for 40 different "noise" alone data sets randomly sampled from the noise distribution described in Section 3.4.2.2. The distribution of the ΔDICs presented in Fig. 3.9 implies the false alarm rate of the burst. Most the ΔDICs $\gtrsim 5$ and the tail of the distribution extends to 18. There are about 10 ΔDICs < 5 and the smallest one is -3.2, which is negative but not small enough to indicate the data favors the positive hypothesis. Therefore, no ΔDICs show that the data strongly favors the positive hypothesis. There are 75% of the ΔDICs implying the data favors the null hypothesis and there are 25% of the ΔDICs implying the data almost equally favors the positive and the null hypothesis.

3.5 Conclusions

In this chapter, we use a Bayesian nonparametric method to analyze the pulsar timing array data set which may contain a contribution from a gravitational wave burst. We have investigated

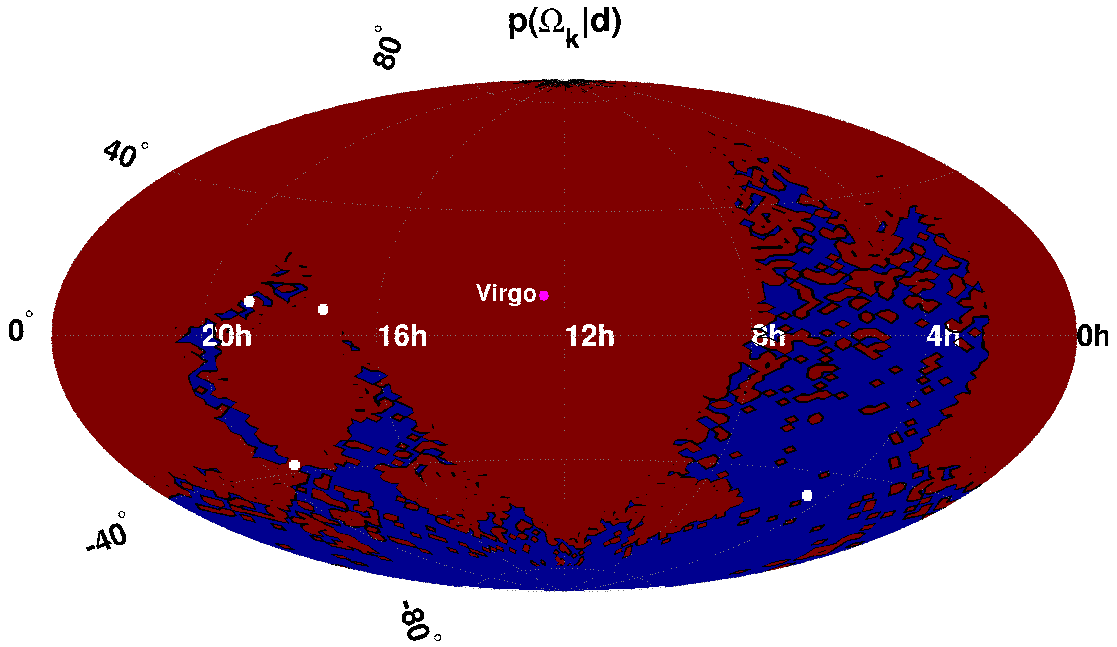


Fig. 3.6 Posterior probability density that the source is found at location Ω_k for the analysis on "weak signal" data set. The inferred sky location has a 99% probability of staying within the red region, and the actual sky location is labelled as "Virgo". The white squares show the locations of the 4 IPTA pulsars used as our pulsar timing array. See main text for details.

how this technique can be used to determine whether a gravitational wave burst is present in the data and, if so, to infer the sky location of the source and the duration of the burst.

Even though we may not know exactly the analytical formulae for the waveform of the bursts, we still have some expectations of the properties of the burst: (1) it should be a smooth function of time; (2) it should have a characteristic duration which should be shorter than the observation duration. By using the Bayesian nonparametrics method, we characterize these two important properties of the gravitational wave bursts into the Gaussian process prior of Eq. (3.12). Correspondingly we are able to detect gravitational wave bursts which are too weak to allow their sources to localized, and make accurate inferences on the sky location of the sources and the shape of the signals when the sources are strong. Compared with the analysis in Finn & Lommen (2010), we infer an important piece of information of the burst, i.e, burst duration. Finn & Lommen (2010) ignored the temporal correlation of the gravitational waveform, and they cannot incorporate the

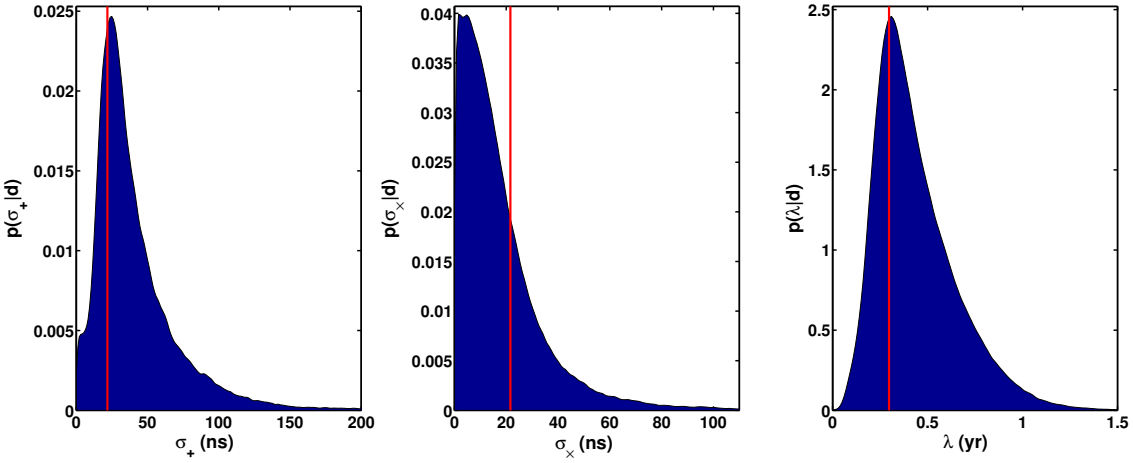


Fig. 3.7 Posterior probability densities of 3 hyperparameters — σ_+, $\sigma_\times$ and λ, for the analysis on the "weak signal" data described in Section 3.4.3.2. The left panel shows that the fractional uncertainty of σ_+ is about 88.1%, and the red line represents the peak gravitational wave induced timing residual of the "+" polarization component of the simulated source. The middle panel shows that the fractional uncertainty of $\sigma_\times$ is beyond 100% and the red line represents the peak gravitational wave induced timing residual of the "×" polarization component of the simulated source. The right panel shows that the fractional uncertainty of λ is about 50%, and the red line represents the duration of the simulated burst divided by $\sqrt{2}$.

important characteristics of the bursts described above into their analysis, which loses the key power of the Bayesian nonparametric methods.

For the purpose of demonstration, we apply our Bayesian nonparametrics analysis to the pulsar timing data of the 4 best millisecond pulsars in the current IPTA, as the capability of detection and characterization of gravitational waves will be dominated by these pulsars (Burt et al., 2011). However, our analysis can be straightforwardly applied to analyze the data of all the pulsars in IPTA. In the future, the effective number of pulsars whose timing noises are low enough to detect gravitational waves is expected to signicantly increase with the commissioning of more sensitive radio telescopes such as the Five-hundred-meter Aperture Spherical Telescope (Nan et al., 2011) and the Square Kilometer Array (SKA) (Dewdney et al., 2009). Applying our analysis method to the pulsar timing data accumulated by these future telescopes will significantly improve the detection sensitivity and inference of the sky location of the sources.

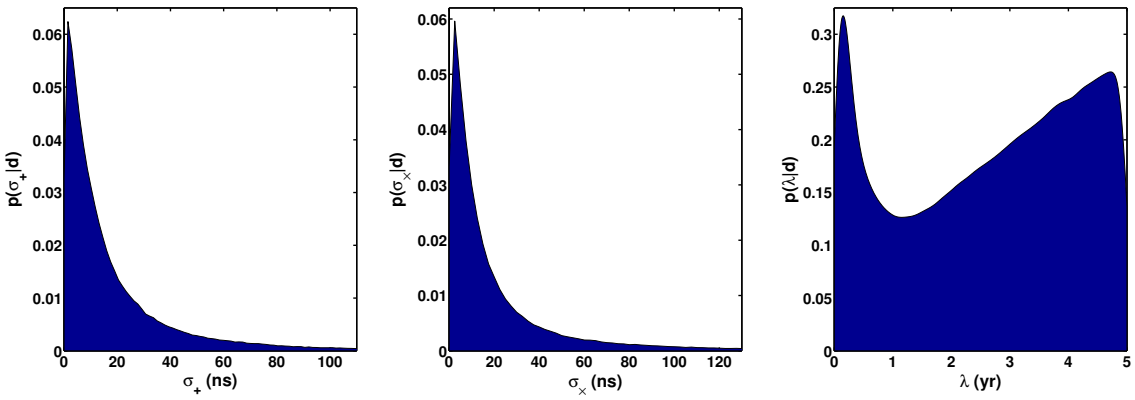


Fig. 3.8 Posterior probability densities of 3 hyperparameters — σ_+, $\sigma_\times$ and λ, for the analysis on the "noise only" data described in Section 3.4.3.3.

While the context of our discussion is gravitational wave burst detection via pulsar timing arrays, following the discussion in Section 3.2 and choosing an appropriate Gaussian process priors with an appropriate kernel, our analysis itself can be directly applied to detecting any kinds of signals whose analytical formulae are unknown, such as in the detection of a binary merger by LIGO (Aasi et al., 2013), the detection of dynamical chaos in exoplanetary systems by the Kepler Mission (Deck et al., 2012), the detection of ultra-high energy gamma rays from gamma ray bursts driven by magnetohydrodynamics by the Fermi satellite (Cenko et al., 2011), etc. In particular, most of these processes can be simulated numerically, which provides ample physical information that can help us choose appropriate kernels, either stationary or non-stationary, smooth or continuous, periodic or chirping, etc. (Rasmussen & Williams, 2006), which makes Bayesian nonparametrics a promising data analysis methodology in physics and astronomy.

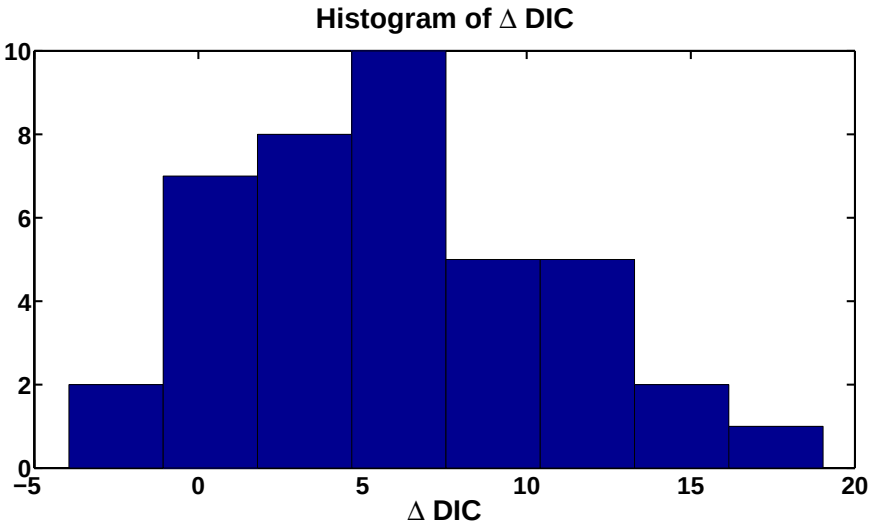


Fig. 3.9 Histogram of the 40 ΔDICs obtained by the Bayesian nonparametric analysis of the burst source. Most the ΔDICs $\gtrsim 5$ and the tail of the distribution extends to 18. There are about 10 ΔDICs < 5 and the smallest one is -3.2. There are 75% of the ΔDICs implying the data favors the null hypothesis and there are 25% of the ΔDICs implying the data almost equally favors the positive and the null hypothesis.

Chapter 4

Search for a Generic Gravitational Wave Background with Bayesian Nonparametric Analysis via Pulsar Timing Arrays

4.1 Introduction

The gravitational wave background results from the superposition of gravitational waves generated from all sources across the universe. The background has been traditionally assumed to be a stochastic Gaussian process with a power law spectrum due to the central limit theorem (e.g., Hellings & Downs, 1983; Jenet et al., 2005, 2006; van Haasteren et al., 2009, 2011). This approximation might break down because the gravitational wave contribution to the pulsar timing signals may be dominated by the strongest sources, the number of which may not be sufficient enough to fulfill the requirement of the central limit theorem (Ravi et al., 2012). Correspondingly previous analyses (e.g., Jenet et al., 2005; van Haasteren et al., 2009) may not have worked effectively. Here we propose a methodology — Bayesian nonparametric analysis (Ghosh & Ramamoorthi, 2003; Hjort et al., 2010) — that can be used to analyze the pulsar timing array data set that potentially includes a contribution from a generic gravitational wave background. This method will set strong constraints on the feasible patterns of the background signals we try to extract from the data. It would investigate whether the gravitational wave background is present and also estimate the parameters that characterize the background. We will see that when the non-Gaussianity of the background becomes non-negligible, our method is still efficacious while the conventional methods become much less effective. Our Bayesian nonparametric analysis is able to

infer whether a gravitational wave background is present in the data, and also estimate the parameters that characterize the background. While the context of our discussion focuses on pulsar timing arrays, the analysis itself is directly applicable to detecting and characterizing any signals that arise from the superposition of a large number of astrophysical events.

In Section 4.2, we discuss the characteristics of the gravitational wave background and its non-Gaussianity. In Section 4.3, we describe a Bayesian nonparametric analysis of the pulsar timing array data. In Section 4.4, we illustrate the effectiveness of this analysis by applying it to several representative examples and compare the analysis with the conventional method assuming the background is Gaussian. Finally, we summarize our conclusion in Section 4.5.

This work is published as X. Deng, 2014, Physical Review D, 90, 104029.

4.2 Characteristics of the gravitational wave background

4.2.1 Pulsar timing residuals induced by a gravitational wave background

The jth pulsar timing response to a gravitational wave background would be the sum of Eq. (1.29),

$$\tau_j(t) = \sum_a \sum_A F_j^{(A)}(\widehat{k}_a) \left[\tau^a_{(A)}(t) - \tau^a_{(A)}(t - L_j(1 + \widehat{k}_a \cdot \hat{n}_j))\right] \tag{4.1}$$

where a labels the contribution from the ath source. We may express the right hand side of Eq. (4.1) in terms of its Fourier components:

$$\begin{aligned}\tau_j(t) &= \sum_a \sum_A F_j^{(A)}(\widehat{k}_a) \int \mathrm{d}f \left[\widetilde{\tau^a_{(A)}}(f) e^{i\,2\pi f t} - \widetilde{\tau^a_{(A)}}(f) e^{i\,2\pi f\left(t - L_j(1+\widehat{k}_a\cdot\hat{n}_j)\right)}\right] \\ &= \sum_A \int \mathrm{d}f \left[\sum_a F_j^{(A)}(\widehat{k}_a)\widetilde{\tau^a_{(A)}}(f) e^{i\,2\pi f t} - \sum_a F_j^{(A)}(\widehat{k}_a)\widetilde{\tau^a_{(A)}}(f) e^{-i\,2\pi f L_j(1+\widehat{k}_a\cdot\hat{n}_j)} e^{i\,2\pi f t}\right]. \end{aligned} \tag{4.2}$$

We notice that compared with the Earth term, the pulsar term has an extra phase $2\pi f L_j(1+\widehat{k}_a\cdot\hat{n}_j)$. In the pulsar timing array waveband, the pulsar distance is much longer than the gravitational wavelength, i.e., $2\pi f L_j \sim 10^4$. Correspondingly, when summing over all the source across the sky, the extra phase ranges from 0 to $\sim 10^4$. The pulsar term vanishes after summing over the timing responses of all the sources with all different pulsar term phases, and the pulsar timing response to a gravitational wave background, i.e., Eq. (4.1), can be simplified as

$$\tau_j(t) = \sum_a \sum_A F_j^{(A)}(\widehat{k}_a)\, \tau^a_{(A)}(t) \tag{4.3}$$

4.2.2 Non-Gaussianity of the gravitational wave background

The assumption of Gaussianity of the gravitational wave background results from the central limit theorem (Feller, 1945). However, the number of gravitational wave sources that can contribute to the pulsar timing array waveband is limited, which may not be sufficient to fulfill the requirement of the central limit theorem (Ravi et al., 2012). Here we use a toy model of gravitational wave source population to illustrate how the degree of non-Gaussianity increases as the number of gravitational wave sources decreases. We refer readers to Ravi et al. (2012) for details.

The degree of non-Gaussianity of a distribution is usually characterized by the skewness and kurtosis (e.g., Joanes & Gill, 1998). The skewness $\widehat{S}_X$ and the kurtosis $\widehat{K}_X$ of a random variable X are respectively defined as (Joanes & Gill, 1998)

$$\widehat{S}_X = \frac{\langle (X - \langle X\rangle)^3\rangle}{\left(\langle (X - \langle X\rangle)^2\rangle\right)^{3/2}} \tag{4.4a}$$

$$\widehat{K}_X = \frac{\langle (X - \langle X\rangle)^4\rangle}{\left(\langle (X - \langle X\rangle)^2\rangle\right)^{2}} - 3 \tag{4.4b}$$

where $\langle\rangle$ denotes the ensemble average. If X is Gaussian distributed, the skewness $\widehat{S}_X$ and the kurtosis $\widehat{K}_X$ are both zero. Negative skewness indicates the left tail of the distribution is heavier than the right one and positive skewness indicates the right tail is heavier than the left one (Joanes & Gill, 1998); negative kurtosis indicates the distribution has lighter tails than a Gaussian distribution and positive kurtosis indicates the distribution has heavier tails than a Gaussian distribution (Joanes & Gill, 1998). If X is the sum of n zero mean identically and independently distributed random variables, i.e.,

$$X = \sum_a^n x_a \tag{4.5a}$$

with

$$\langle x \rangle = 0 \tag{4.5b}$$

then we can obtain

$$\langle X \rangle = 0 \tag{4.5c}$$

$$\langle (X - \langle X \rangle)^2 \rangle = n \langle x_a^2 \rangle \tag{4.5d}$$

$$\langle (X - \langle X \rangle)^3 \rangle = n \langle x_a^3 \rangle \tag{4.5e}$$

$$\langle (X - \langle X \rangle)^4 \rangle = n \langle x_a^4 \rangle + 3n(n-1)(\langle x_a^2 \rangle)^2 \tag{4.5f}$$

where we have used the fact that $\langle x_a \rangle = 0$. Correspondingly, when n is large, we can express the skewness and kurtosis as

$$\widehat{S}_X = \frac{1}{\sqrt{n}} \frac{\langle x_a^3 \rangle}{\left(\langle x_a^2 \rangle \right)^{3/2}} \tag{4.5g}$$

$$\widehat{K}_X = \frac{1}{n} \frac{\langle x_a^4 \rangle}{\left(\langle x_a^2 \rangle \right)^2}. \tag{4.5h}$$

We can see that as the number of the individual random variable x_a goes to infinity, the skewness and kurtosis approach zero and the distribution becomes Gaussian. This is what the central limit theorem implies.

In the case of the gravitational wave background, X is the timing residual τ_j induced by the background in Eq. (4.3), and the x_a are the timing residuals induced by the ath single source. For simplicity, we assume that the supermassive black hole binaries are the primary sources of gravitational waves in the pulsar timing array waveband. They are homogeneously and isotropically distributed in the sky, and their orbital evolution is driven by gravitational wave emission; correspondingly, the number of the binaries per frequency per comoving volume is (Sesana et al., 2008)

$$\frac{\mathrm{d}^2 N}{\mathrm{d}f\,\mathrm{d}V_\mathrm{c}} \propto f^{-11/3} \tag{4.6}$$

where N denotes the number of the binaries; f and V_c respectively denote gravitational wave frequency and comoving volume of the universe. We also assume that the gravitational waves generated by these sources are monochromatic waves. We only consider gravitational waves with period $\lesssim 5$yr, of which the induced peak pulsar timing residuals are above 0.01 ns, since only these waves will significantly contribute to the pulsar timing residuals across 5 years of observation as in the current IPTA (Hobbs et al., 2010; Demorest et al., 2012; Manchester et al., 2013). The gravitational waves with longer periods will be fitted out by the procedure in the standard pulsar timing analysis that removes the linear and quadratic trends induced by pulsar spin and spin down. In this range, the number of sources N should be $\sim 10^5 - 10^6$ as expected from theoretical models (Sesana et al., 2008, 2011).

To compute the skewness and kurtosis of timing residuals induced by a gravitational wave background, we can directly follow Eq. (4.5g) and Eq. (4.5h). The timing residual τ_j^a of the jth pulsar induced by the sinusoidal gravitational waves generated from the ath individual source would

be

$$\tau_j^a(t) = F_j^{(+)}(\widehat{k}_a)\, B^a_{(+)} \cos\left(2\pi f^a t + \phi^a\right) + F_j^{(\times)}(\widehat{k}_a)\, B^a_{(\times)} \sin\left(2\pi f^a t + \phi^a\right)$$

$$= B_j^a \cos\left(2\pi f^a t + \phi^a + \psi_j^a\right) \tag{4.7a}$$

where $B^a_{(A)}$ denotes the amplitudes of the timing residuals induced by the (A) polarization component; f^a and ϕ^a are respectively the frequency and initial phase of the gravitational waves; B_j^a and ψ_j^a are respectively

$$B_j^a = \sqrt{[F_j^{(+)}\, B^a_{(+)}]^2 + [F_j^{(\times)}\, B^a_{(\times)}]^2} \tag{4.7b}$$

$$\psi_j^a = \arctan\left[\frac{F_j^{(+)}\, B^a_{(+)}}{F_j^{(\times)}\, B^a_{(\times)}}\right]. \tag{4.7c}$$

Assuming that the initial phase ϕ^a is uniformly distributed between 0 and 2π, we can obtain

$$\langle \tau_j^a \rangle = \langle B_j^a \rangle \langle \cos\left(2\pi f^a t + \phi^a + \psi_j^a\right) \rangle = 0 \tag{4.8a}$$

$$\langle \left(\tau_j^a\right)^2 \rangle = \langle \left(B_j^a\right)^2 \rangle \langle \cos^2\left(2\pi f^a t + \phi^a + \psi_j^a\right) \rangle = \pi \langle \left(B_j^a\right)^2 \rangle \tag{4.8b}$$

$$\langle \left(\tau_j^a\right)^3 \rangle = \langle \left(B_j^a\right)^3 \rangle \langle \cos^3\left(2\pi f^a t + \phi^a + \psi_j^a\right) \rangle = 0 \tag{4.8c}$$

$$\langle \left(\tau_j^a\right)^4 \rangle = \langle \left(B_j^a\right)^4 \rangle \langle \cos^4\left(2\pi f^a t + \phi^a + \psi_j^a\right) \rangle = \frac{3\pi}{4} \langle \left(B_j^a\right)^4 \rangle \tag{4.8d}$$

and following Eq. (4.5g) and Eq. (4.5h), the skewness and kurtosis of timing residuals induced by a gravitational wave background would be

$$\widehat{S}_{\tau} = 0 \tag{4.8e}$$

$$\widehat{K}_{\tau} = \frac{1}{n}\frac{3}{4\pi}\frac{\langle\left(B_j^a\right)^4\rangle}{\left[\langle(B_j^a)^2\rangle\right]^2}. \tag{4.8f}$$

The amplitude B_j^a depends on the pattern function, gravitational wave amplitudes and frequency. We can compute $\langle\left(B_j^a\right)^2\rangle$ and $\langle\left(B_j^a\right)^4\rangle$ by sampling an ensemble of the amplitude B_j^a from the distribution Eq. (4.6) and numerically computing the ensemble average. Table 4.1 presents the skewness and kurtosis of timing residuals of PSR J1713+0747 induced by gravitational wave backgrounds respectively generated from 10^6, 5×10^5, 2×10^5 and 10^5 sources. As we stated above, all of these sources are generated in the pulsar timing array waveband, i.e., their gravitational wave periods range from 3 months to 5 years. Their frequency distribution follows Eq. (4.6). For all of these four gravitational wave backgrounds, about 0.1% of the sources are responsible for 95% of the residual power.

Table 4.1 Skewness and kurtosis of timing residuals induced by gravitational wave backgrounds respectively generated from 10^6, 5×10^5, 2×10^5 and 10^5 sources.

No. of Sources	skewness	kurtosis
10^6	0	0.01
5×10^5	0	0.02
2×10^5	0	0.05
10^5	0	0.1

Note. We can see that the kurtosis increases as the number of souces of the gravitational wave background decreases. How large the kurtosis should be to make the non-Gaussianity significant depends on the signal-to-noise ratio of the background, the source distribution of the background, etc. In Section 4.4 we use detailed examples to show that for the source distribution described in this section, when the degree of the non-Gaussianity implied by the kurtosis is significant.

We can see that as the number of sources decreases, the degree of non-Gaussianity increases as implied by the central limit theorem. Correspondingly, the power law spectrum of the gravitational wave background as generally assumed would be effectively modified as it is derived based on the assumption that the number of sources is approximately infinite (e.g., Flanagan, 1993; Jenet et al., 2005; van Haasteren et al., 2009).

4.3 Bayesian nonparametric analysis

To search for evidence of a gravitational wave background in the data set, we need to model its contribution to pulsar timing residuals. In order to avoid the strong assumption of Gaussianity, we introduce a different method — Bayesian nonparametric analysis (Ghosh & Ramamoorthi, 2003; Rasmussen & Williams, 2006; Hjort et al., 2010). This method treats the gravitational wave background as a *deterministic* process since it originates from the superposition of gravitational waves from a finite number of sources, each of which is a deterministic process. However, we do not try to write down the exact deterministic function form. Instead, we assign a prior distribution on the function form of the signal, which will characterize the expected shape of the signal pattern, such as its smoothness, variation, trend, etc. These characteristics are used to represent the function form of the signal. Correspondingly, we are able to detect the signal whose deterministic function form has the same characteristics as what our prior distribution characterizes. For discussion of application of Bayesian nonparametrics in the gravitational wave context, see (Deng, 2014; Lentati et al., 2014; Lee et al., 2014) for details.

4.3.1 The choice of Gaussian process priors

4.3.1.1 Prior of $\boldsymbol{\tau}$

For the ath source, we choose the same kernel of the gaussian process prior as in Section 3.3.2.1:

$$\mathrm{K}_{a(A)}(\Delta t) = \sigma^2_{a(A)} \exp(-\frac{\Delta t^2}{2\,\lambda_a^2}) \tag{4.9a}$$

and the corresponding Gaussian process prior of $\tau^a_{(A)}$ would be

$$q(\tau^a_{(A)}) = \frac{\exp\left[-\frac{1}{2}\int \mathrm{d}t\,\mathrm{d}t'\,\tau^a_{(A)}(t)\,\mathrm{K}^{-1}_{a(A)}(\Delta t)\,\tau^a_{(A)}(t')\right]}{\sqrt{(2\pi)^{\mathrm{dimK}}\det||\mathrm{K}||}} \tag{4.9b}$$

where $\sigma_{a(A)}$ represents the rms amplitude $\tau^a_{(A)}$ from the ath source and λ_a is the characteristic time-scale of the waveform. Within λ_a, $\tau^a_{(A)}$ is expected to cross the zero level only once (Adler, 1981). Therefore, for gravitational wave burst sources such as encounters of two supermassive black holes, λ_a characterizes the duration of the burst (Deng, 2014); and for gravitational waves from gravity-bound binaries, λ_a characterizes periods of the binaries.

We denote the timing residuals induced by the superposition of gravitational waves from the direction $\hat{k}$ as $\tau_{\hat{k}}$. Eq. (4.3) shows that $\tau_{\hat{k}}$ is the linear superposition of $\tau^a_{(A)}$ and we assume that two polarization components of the gravitational wave from a single source are independent; correspondingly, the kernel for the Gaussian process prior of $\tau_{\hat{k}}$ would be the linear superposition of Eq. (4.9a),

$$\mathrm{K}_{\hat{k}\,(jk)}(\Delta t) = \sum_A F^{(A)}_j(\hat{k}) F^{(A)}_k(\hat{k}) \sum_{a\,(\hat{k})} \sigma^2_{a(A)} \exp\left(-\frac{\Delta t^2}{2\,\lambda_a^2}\right) \tag{4.10a}$$

where j, k denote pulsar indices and $a\,(\hat{k})$ denotes that the sum is over all sources in sky location $\hat{k}$. If we assume that at the sky location $\hat{k}$, the total square sum of all source rms amplitudes $\sigma_{a(A)}$ is $\sigma^2_{\hat{k}(A)}$ and the density of the sources with square characteristic time-scale $\lambda^2_{\hat{k}}$ is $\pi_{\hat{k}(A)}\left(\lambda^2_k\right)$, we can approximate the second sum in Eq. (4.10a) as an integral over all possible $\lambda^2_{\hat{k}}$, i.e.,

$$\sum_{a\,(\hat{k})} \sigma^2_{a(A)} \exp\left(-\frac{\Delta t^2}{2\,\lambda^2_a}\right) = \sigma^2_{\hat{k}(A)} \int \mathrm{d}\lambda^2_{\hat{k}}\; \pi_{\hat{k}(A)}(\lambda^2_{\hat{k}}) \exp\left(-\frac{\Delta t^2}{2\,\lambda^2_{\hat{k}}}\right). \tag{4.10b}$$

Because the distribution of polarization angles is expected to be uniform (Misner et al., 1973; Christensen, 1992; Flanagan, 1993), so $\sigma^2_{\hat{k}(A)}$ and $\pi_{\hat{k}(A)}$ are the same for the two GW polarization components, i.e.,

$$\sigma^2_{\hat{k}(+)} = \sigma^2_{\hat{k}(\times)} = \sigma^2_{\hat{k}} \tag{4.10c}$$

$$\pi_{\hat{k}(+)} = \pi_{\hat{k}(\times)} = \pi_{\hat{k}}. \tag{4.10d}$$

As a result, the kernel of the Gaussian process prior of $\tau_{\hat{k}}$ becomes

$$\mathrm{K}_{\hat{k}\,(jk)}(\Delta t) = \left(\sum_A F^{(A)}_j(\hat{k}) F^{(A)}_k(\hat{k})\right) \sigma^2_{\hat{k}} \int \mathrm{d}\lambda^2_{\hat{k}}\; \pi_{\hat{k}}(\lambda^2_{\hat{k}}) \exp\left(-\frac{\Delta t^2}{2\,\lambda^2_{\hat{k}}}\right). \tag{4.10e}$$

We choose $\pi_{\hat{k}}$ as an inverse gamma distribution of the square characteristic time-scale (Stein, 1999; Rasmussen & Williams, 2006), which is conjugate to the square exponential $\exp\left(-\frac{\Delta t^2}{2\,\lambda^2_{\hat{k}}}\right)$,

$$\pi_{\hat{k}}(\lambda^2_{\hat{k}}) = \frac{\beta^{\alpha_{\hat{k}}}_{\hat{k}}}{\Gamma(\alpha_{\hat{k}})} \lambda^{-2\alpha_{\hat{k}}-2}_{\hat{k}} \exp\left(-\frac{\beta_{\hat{k}}}{\lambda^2_{\hat{k}}}\right) \tag{4.11a}$$

$$\text{or} \quad \pi^{1/\lambda}_{\hat{k}}(1/\lambda^2_{\hat{k}}) = \frac{\beta^{\alpha_{\hat{k}}}_{\hat{k}}}{\Gamma(\alpha_{\hat{k}})} (1/\lambda_{\hat{k}})^{2\alpha_{\hat{k}}-2} \exp\left(-\frac{\beta_{\hat{k}}}{\lambda^2_{\hat{k}}}\right) \tag{4.11b}$$

where $\alpha_{\hat{k}}$ and $\beta_{\hat{k}}$ are respectively referred to as the shape parameter and the scale parameter, and they both have to be positive to guarantee that $\pi_{\hat{k}}$ is normalizable (Bernardo & Smith, 2003; Gelman et al., 2004); $\Gamma(\alpha_{\hat{k}})$ is the gamma function of $\alpha_{\hat{k}}$ (Abramowitz & Stegun, 1964). Combining Eq. (4.10e) and Eq. (4.11), we can obtain the kernel of the Gaussian process prior of $\tau_{\hat{k}}$

$$\mathrm{K}_{\hat{k}\,(jk)}(\Delta t) = \sigma_{\hat{k}}^2 \left(1 + \frac{\Delta t^2}{2\alpha_{\hat{k}}\xi_{\hat{k}}^2}\right)^{-\alpha_{\hat{k}}} \sum_A F_j^{(A)}(\hat{k}) F_k^{(A)}(\hat{k}) \tag{4.12}$$

where $\xi_{\hat{k}} = \sqrt{\beta_{\hat{k}}}$, which is the characteristic time-scale of $\tau_{\hat{k}}$.

Because the gravitational wave background is the superposition of the gravitational waves from all directions in the sky, the kernel of the Gaussian process prior of the timing residuals induced by the background should be the sum of Eq. (4.12) across the whole sky. In general, the gravitational wave sources may not be isotropically distributed and the anisotropy of the background can be characterized by decomposing the angular distribution of the gravitational wave energy density on the sky into multipole moments (Mingarelli et al., 2013; Taylor & Gair, 2013). However, for the purpose of demonstration, in this chapter we only focus on the isotropic gravitational wave background and it is straightforward to generalize our method to anisotropic cases by combining the techniques presented in (Mingarelli et al., 2013; Taylor & Gair, 2013). By assuming isotropy, $\tau_{\hat{k}}$ in all directions will have the same $\sigma_{\hat{k}}$, $\alpha_{\hat{k}}$ and $\xi_{\hat{k}}$. Correspondingly, the kernel of the gravitational wave background induced timing residuals $\boldsymbol{\tau}$ is

$$\mathrm{K}_{jk}(\Delta t) = \sigma^2 \left(1 + \frac{\Delta t^2}{2\alpha\xi^2}\right)^{-\alpha} \gamma_{jk} \tag{4.13a}$$

where γ_{jk} is the Hellings-Downs curve (Hellings & Downs, 1983).

$$\begin{aligned}\gamma_{jk} &= \int \mathrm{d}^2\Omega_{\hat{k}} \sum_A F_j^{(A)}(\hat{k}) F_k^{(A)}(\hat{k}) \\ &= \frac{3}{2}\frac{1-\hat{n}_j\cdot\hat{n}_k}{2}\log\left(\frac{1-\hat{n}_j\cdot\hat{n}_k}{2}\right) - \frac{1}{4}\frac{1-\hat{n}_j\cdot\hat{n}_k}{2} + \frac{1}{2} + \frac{1}{2}\delta_{jk}. \end{aligned} \tag{4.13b}$$

Using Eq. (4.13), we can obtain the Gaussian process prior of $\boldsymbol{\tau}$

$$q(\boldsymbol{\tau}|\sigma,\xi,\alpha) = \frac{\exp\left[-\frac{1}{2}\boldsymbol{\tau}^T\mathbf{K}^{-1}\boldsymbol{\tau}\right]}{\sqrt{(2\pi)^{\dim\mathbf{K}}\det||\mathbf{K}||}} \tag{4.14}$$

where $\mathbf{K}$ is expressed by Eq. (4.13).

4.3.1.2 Prior of hyperparameters

The prior probability density of $\boldsymbol{\tau}$ Eq. (4.14) includes three undetermined hyperparameters — σ, ξ and α, and we need to choose their prior probability density to make a full Bayesian inference.

For σ, this is a scale factor and we can choose the Jeffreys prior (Jeffreys, 1946). However, the Jeffreys prior will make the posterior probability density improper (Gelman, 2006) and correspondingly such an uninformative prior is not an appropriate one. In this case, a uniform distribution from 0 to $+\infty$, which will make the posterior distribution normalizable, is recommended (Gelman, 2006), i.e.,

$$q_\sigma(\sigma) \propto 1. \tag{4.15a}$$

For ξ, this is a time-scale factor and we choose the Jeffreys prior

$$q_\xi(\xi) \propto 1/\xi. \tag{4.15b}$$

The hyperparameter α is the shape parameter of the inverse gamma distribution Eq. (4.11), which represents the number density of the sources. We expect that the number of sources should monotonically decrease with their periods or durations (Sesana et al., 2008), and correspondingly, $\pi_{\hat{k}}$ in Eq. (4.11) should be a monotonically increasing function of $\lambda_{\hat{k}}$. To satisfy this requirement, α should be between 0 and 1 (see Eq. (4.11)). We choose the prior distribution for α as a uniform distribution from 0 to 1, i.e.,

$$q_{\alpha}(\alpha) = 1. \tag{4.15c}$$

Since the characteristic timescale λ represents the period or duration of the individual source, so $1/\lambda$ represents the characteristic frequency of the source. Given the prior for α in Eq. (4.15c), we can obtain the prior distribution for the characteristic frequency of the single source marginalizing over α by following Eq. (4.11b):

$$q(1/\lambda^2|\beta) = \int \pi^{\lambda}(1/\lambda^2|\alpha,\beta)q(\alpha)\mathrm{d}\alpha \tag{4.16}$$

where β represents the average square period of the sources over the sky as discussed above. Fig. 4.1 plots q with different βs.

4.3.2 Discussion of deterministic and stochastic modeling

At this point, it is worth comparing the deterministic modeling by Bayesian nonparametrics discussed above and the stochastic modeling used by previous methods.

In general, when we detect a signal across a finite time duration, we have two ways to model the signal:

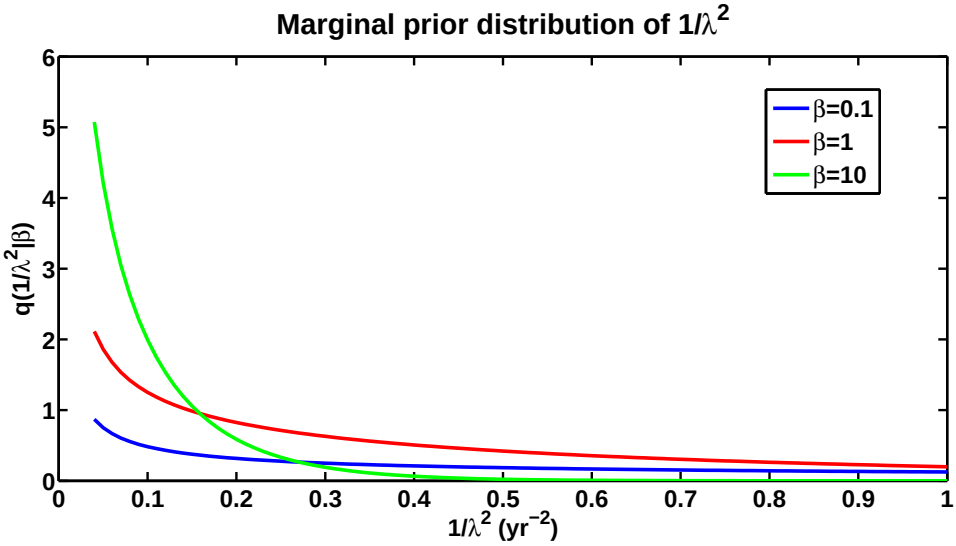


Fig. 4.1 Marginal prior disribution for the individual source's square characteristic frequency $1/\lambda$ with different averaging square period β, where β is in the units of yr^2. The prior sets a constraint that the number of sources monotonically declines with the source frequencies.

(1) Assume the signal is generated by a deterministic process. However, we do not know the function form of the signal, on which we need to assign a prior distribution. This is what we did in this chapter.

(2) Assume the signal is a random sample (one single realization) generated from a stochastic process, and what we need to do is to model the distribution function of the stochastic process. This is what previous methods did.

Because we do not have an ensemble of the signals, both of these modeling methods may lead to reasonable characterizations of the signal. Which is more effective depends on which model fits the data better, i.e., which method leads to a larger likelihood or a smaller DIC.

In the case of detecting a gravitational wave background, both of these methods would result in the same marginal likelihood function Eq. (3.15d), because in the first method, we assume that the prior distribution of the signal form is a Gaussian distribution, and in the second method,

we assume that the distribution function of the gravitational wave background is also a Gaussian distribution. This means that both of these two methods may lead to the same inference of the signal. So in practice, the method we present above may be considered as the same method as previous ones in (van Haasteren et al., 2009; Lentati et al., 2013) except for the difference choices of the kernel $\mathbf{K}$. However, the choices of the kernels strongly depend on what logic we follow, which will lead to different values of likelihood functions. This is the key difference between the method presented here and the previous methods.

When we follow the first approach, as described in Section 4.3.1.1, the kernel $\mathbf{K}$, which originally appears in the prior distribution, is chosen to characterize the expected shape of the signal, such as its smoothness, its variation, its trend, etc. We use these characteristics to represent the underlying unknown deterministic function form of the signal. Following this logic, we finally obtain the appropriate kernel Eq. (4.13). Meanwhile, if we follow the second approach, $\mathbf{K}$ characterizes the covariance of the Gaussian distribution, which is assumed to be the underlying distribution function of the gravitational wave background. Following this logic, $\mathbf{K}$ would have to be the Fourier transform of a frequency power law, because the frequency distribution of the gravitational wave sources follows a power law (van Haasteren et al., 2009; Lentati et al., 2013). Using the first method cannot lead to the choice of a power law spectrum while using the second method cannot lead to the choice of Eq. (4.13). The two different choices of $\mathbf{K}$ would lead to different values of the likelihood functions.

If the non-Gaussian part of the gravitational wave background is significant, the second method will be ineffective because it only characterizes the covariance of the distribution but ignores the skewness, kurtosis and other parts of the distribution. The first method, however, assumes the signal is just a representation of some deterministic function but not a random sample drawn from some underlying distribution. We assign a prior distribution with a specific kernel to

characterize the expected shape of the signal. Correspondingly, no matter whether the signal is sampled from some distribution or what the distribution is, as long as our Gaussian prior with a kernel correctly characterizes the shape of the signal pattern, such as its smoothness, variation, trend, etc., our method will lead to a good inference of the signal.

To recap, although the deterministic modeling by our Bayesian nonparametrics and the previous stochastic modeling will both lead to likelihood functions with the same form, our method only tries to characterize the signal itself and the kernel in the Gaussian prior distribution represents our expectation of the shape of the signal pattern. Previous methods, on the other hand, assume that the signal is sampled from some underlying distribution and try to use a Gaussian model with a power-law spectrum to characterize the distribution. These two approaches will lead to different choices of the kernels, which correspondingly results in different values of likelihood functions.

4.4 Examples

4.4.1 Overview

To illustrate the effectiveness of the analysis techniques described above, we apply them to simulated data sets of the pulsar timing array described in Table 3.1.

We uniformly sample 50 observation times across 5 years of observation, and the corresponding pulsar timing data sets are constructed by

(1) Evaluating the pulsar timing residuals induced by a simulated gravitational wave background that will be described in Section 4.4.2.

(2) Adding pulsar timing noise described in Section 3.4.2.2 to the timing residuals obtained by the first step.

(3) Removing the linear trend of the timing residuals obtained above to simulate the procedure in the standard pulsar timing analysis that removes the effects of pulsar spin and spin down.

When we analyze the data, we add a linear model in the pulsar timing response function to account for the linear trend, the same as the analysis in van Haasteren et al. (2009). We will also apply our analysis methods to a data set composed of timing noise alone for comparative study.

4.4.2 Construction of simulated signals of the gravitational wave background

We construct the simulated observations of two isotropic gravitational wave backgrounds respectively generated from 10^6 and 10^5 supermassive black hole binaries. Both of these sources are generated in the same way as we did in Section 4.2.2:

(1) The frequency distribution of these sources follows Eq. (4.6) with lower bound of $0.2\,\mathrm{yr}^{-1}$ and upper bound of $4\,\mathrm{yr}^{-1}$.

(2) All of these sources are uniformly distributed across the sky with their gravitational wave peak timing residuals ranging from 0.01ns to 100ns.

(3) The orbital orientations and initial phases of all these sources are uniformly distributed.

(4) The timing residuals induced by the gravitational wave background are the sum of all the gravitational wave signals generated from the sources sampled from the distribution described in the above three steps. The rms values of the gravitational wave amplitudes of the two gravitational wave backgrounds are both about 22ns.

The degrees of non-Gaussianity of these two backgrounds are presented in Table 4.1, and we can see that the background with 10^5 sources is more non-Gaussian than that with 10^6 sources. For these two cases of simulated backgrounds, we will compare the results of our method and the

conventional one proposed by van Haasteren et al. (2009) and we will see that our method is much more effective in the case of the background with 10^5 sources.

4.4.3 Analysis of simulated data sets

Our Bayesian nonparametric analysis is designed to investigate whether a gravitational wave background is present in the dataset, and also infer the hyperparameters.

We apply both our method and the conventional Gaussian method on the two data sets — one contains the contribution of the gravitational wave background with 10^6 sources and the other contains the contribution of the background with 10^5 sources. We also analyze the data set consisting of timing noise alone by our method for comparative study. Table 4.2 lists the results of the analysis. The parentheses in the second column contain the DIC differences obtained by applying the conventional Gaussian method in van Haasteren et al. (2009).

Table 4.2 Summary statistics of the gravitational wave background analysis. The symbol ϵ_{σ}, ϵ_{ξ} and ϵ_{α} respectively denote the fractional uncertainties of σ, ξ and α. The parentheses contain the DIC differences by applying the conventional Gaussian method on the same data sets. The convergence diagnostics show our MCMC samples are convergent.

No. of Sources	ΔDIC (Gaussian)	ϵ_{σ}	ϵ_{ξ}	ϵ_{α}	max Z-score
10^6	-15 (-12)	27.4%	96.8%	32.0%	0.210
10^5	-14 (-3)	29.2%	71.9%	34.2%	0.308
Absent	5 (5)	92.5%	63.5%	54.6%	0.218

4.4.3.1 Signal of an Isotropic Gravitational Wave Background with 10^6 Sources

We simulate an isotropic gravitational wave background by sampling 10^6 sources from a homogeneous and isotropic distribution Eq. (4.6) and computing the superposition of the timing residuals induced by the gravitational waves from them, as described in Section 4.4.2. The first row of Table. 4.2 and Fig. 4.2 summarize the results of our Bayesian nonparametric analysis:

(1) From the first row of Table 4.2, we see that the difference between the DICs of the positive hypothesis and null hypothesis is -15, corresponding to a strong evidence for the presence of a gravitational wave background in the data set. We apply the conventional Gaussian method in van Haasteren et al. (2009) on the same data set and the DIC difference is -12, which also indicates a strong evidence of the presence of a gravitational wave background. Therefore, the non-Gaussianity of this background is not large enough to distinguish the effectiveness of the two analysis methods.

(2) We also infer the hyperparameters and Fig. 4.2 shows their posterior probability density. The mean value of σ is 22.5 ns (consistent with the simulated signal of rms amplitude 22ns) and its rms uncertainty is 6.2 ns. Correspondingly, its fractional uncertainty is 27.4%. For ξ, the mean value is 0.25 yr and the rms uncertainty is 0.24 yr, and the corresponding fractional uncertainty is 96.8%. We cannot measure the shape parameter α very well and it tends to be 1. The mean and rms uncertainty are respectively 0.69 and 0.22, and the fractional uncertainty is 32.0%. Fig. 4.3 shows the posterior probability density of the strain amplitude $\sigma_{\rm GW}$ and the spectrum power index $\alpha_{\rm GW}$ of the gravitational wave background, obtained by assuming that the background is Gaussian and applying the conventional method in van Haasteren et al. (2009).

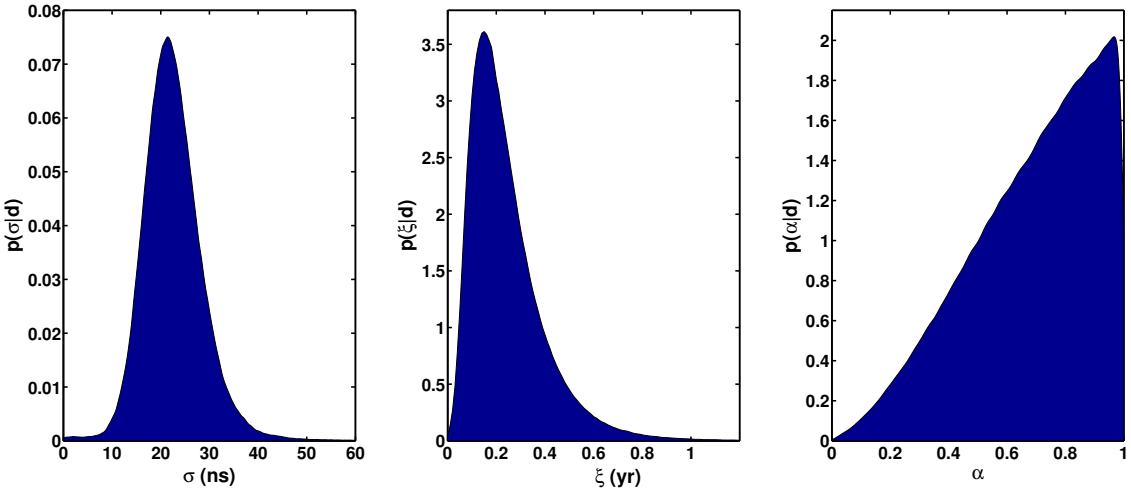


Fig. 4.2 Posterior probability densities of 3 hyperparameters — σ, ξ and α, for analysis on the data described in Section 4.4.3.1. Their fractional uncertainties are respectively 27.4%, 96.8%, 32.0%.

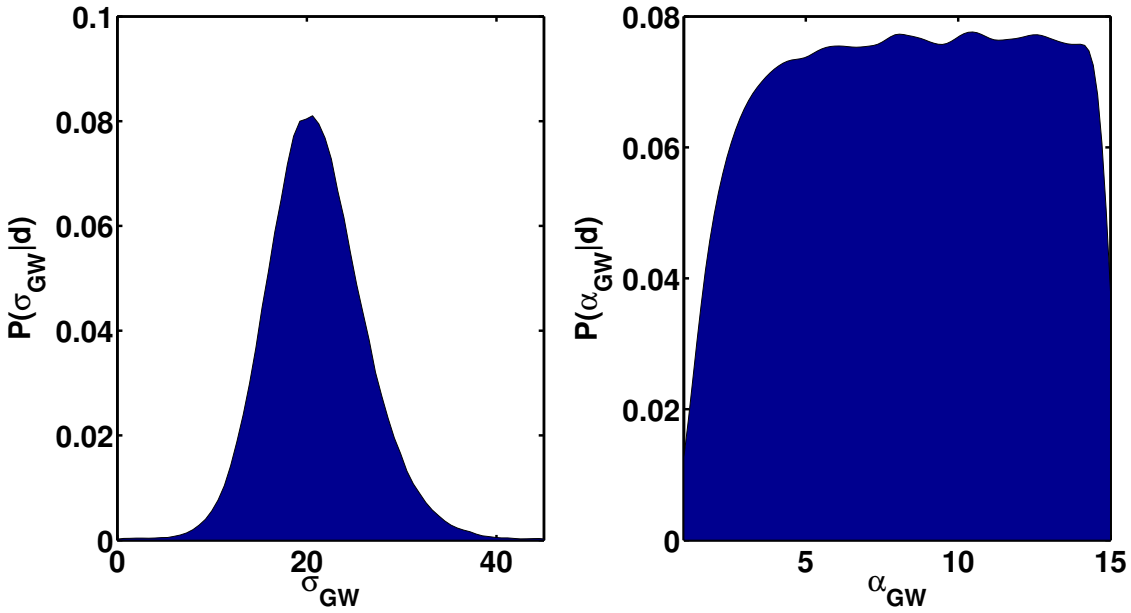


Fig. 4.3 Posterior probability densities of the strain amplitude $\sigma_{\rm GW}$ and the spectrum power index $\alpha_{\rm GW}$ of the gravitational wave background with 10^6 sources obtained by assuming that the background is Gaussian and applying the conventional method in van Haasteren et al. (2009).

4.4.3.2 Signal of an Isotropic Gravitational Wave Background with 10^5 Sources

We simulate an isotropic gravitational wave background by sampling 10^5 sources from a homogeneous and isotropic distribution Eq. (4.6) and computing the superposition of the timing residuals induced by the gravitational waves from them, as described in Section 4.4.2. The degree of non-Gaussianity of this background is greater than the one with 10^6 sources as illustrated in Table 4.1. The second row of Table 4.2 and Fig. 4.4 summarize the results of our Bayesian nonparametric analysis on such "anisotropic signal" data:

(1) From the second row of Table 4.2, we see that the difference between the DICs of the positive hypothesis and null hypothesis is -14, corresponding to strong evidence for the presence of a gravitational wave background in the data set. When we apply the conventional Gaussian method in van Haasteren et al. (2009) on the data set, we obtain a DIC difference of only -3, which indicates no strong evidence of a gravitational wave background. Therefore, the non-Gaussianity of this background is non-negligible, and our method shows a strong advantage over the conventional one in this case.

(2) We also infer the hyperparameters and Fig. 4.4 shows their posterior probability densities. The mean value of σ is 22.4 ns (consistent with the simulated signal of rms amplitude 22ns) and its rms uncertainty is 6.5 ns. Correspondingly, its fractional uncertainty is 30.0%. For ξ, the mean value is 0.36 yr and the rms uncertainty is 0.26 yr, and the corresponding fractional uncertainty is 71.9%. We cannot measure the shape parameter α very well either and it also tends to be 1. The mean and rms uncertainty are respectively 0.67 and 0.23, and the fractional uncertainty is 34.2%. Fig. 4.5 shows the posterior probability density of the strain amplitude $\sigma_{\rm GW}$ and the spectrum power index $\alpha_{\rm GW}$ of the gravitational wave background

obtained by assuming that the background is Gaussian and applying the conventional method in van Haasteren et al. (2009).

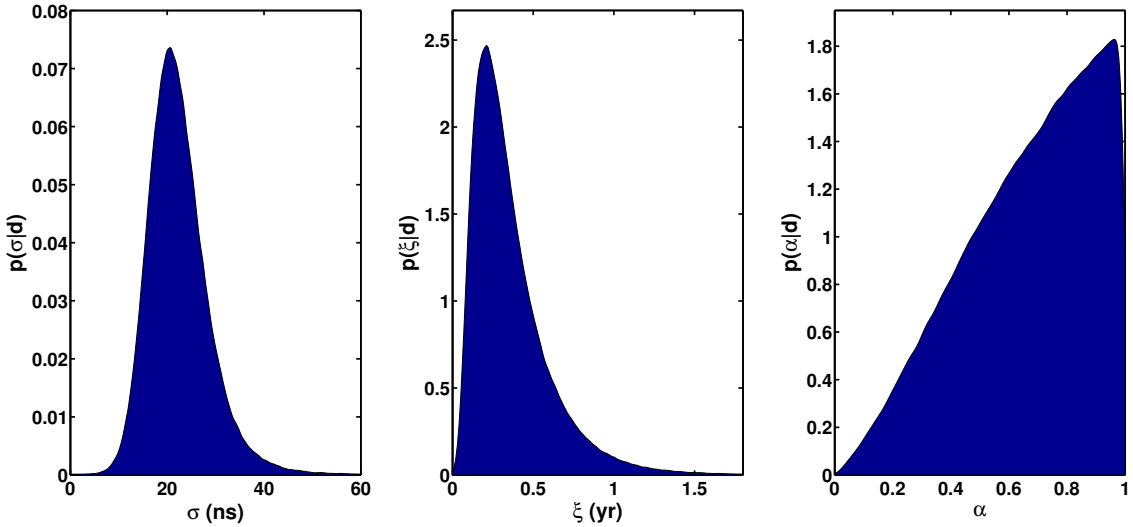


Fig. 4.4 Posterior probability densities of 3 hyperparameters — σ, ξ and α, for analysis on the data described in Section 4.4.3.2. Their fractional uncertainties are respectively 30.0%, 71.9%, 34.2%.

4.4.3.3 No Signal

For comparative study, we also apply our Bayesian nonparametric analysis to a data set with timing noise alone. The third row of Table 4.2 and Fig. 4.6 summarize the results. The difference between the DICs of the two repulsive hypotheses is 5, which shows that the data favors the null hypothesis. All the hyperparameters are imprecisely determined. The Gaussian method in van Haasteren et al. (2009) also offers a DIC difference of 5.

4.4.4 Analysis of IPTA mock data challenge

I test the Bayesian nonparametric method on the closed data set 2 in the IPTA mock data challenge[1] and compare the result with the previous work (e.g., Ellis et al., 2012a). The noise

[1] url: http://www.ipta4gw.org/page_id=116

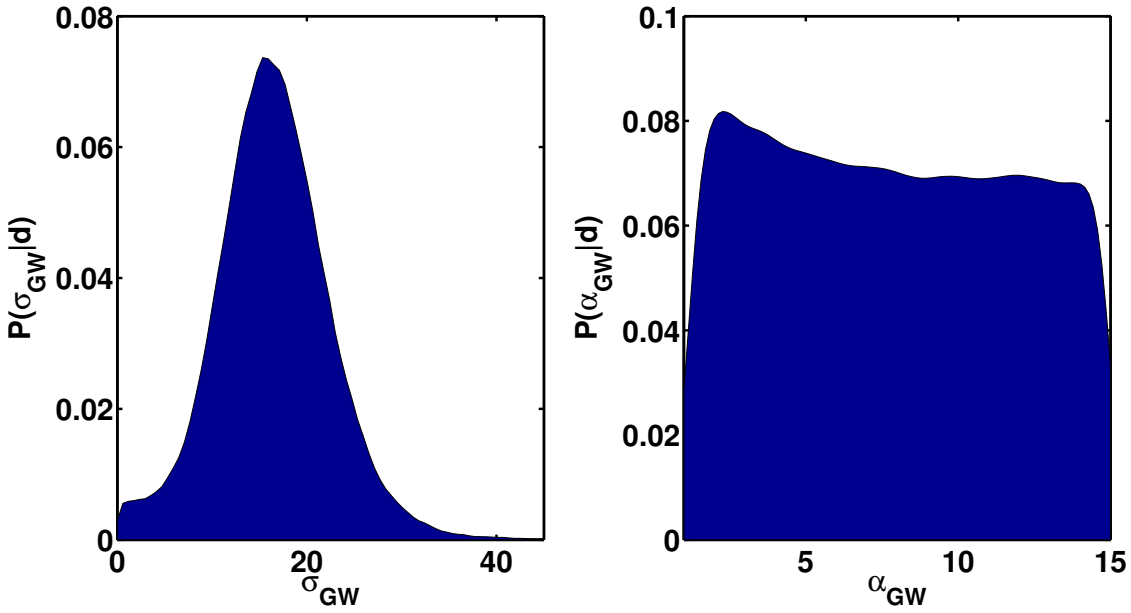


Fig. 4.5 Posterior probability densities of the strain amplitude $\sigma_{\rm GW}$ and the spectrum power index $\alpha_{\rm GW}$ of the gravitational wave background with 10^5 sources obtained by assuming that the background is Gaussian and applying the conventional method in van Haasteren et al. (2009).

in the data is white noise with different rms errors for different pulsars. The signal is a random sample generated from a Gaussian process with a strain amplitude of 6×10^{-14} at a frequency of $1\,{\rm yr}^{-1}$ and a power law spectrum with index of $-2/3$. The observation times are unifomly sampled across 5 years with an averaged observational interval of two weeks. The amplitude signal-to-noise ratio for this source is ~ 100.

The DIC difference of this data set obtained by the Bayesian nonparametric method is -600 (approximately a log Bayes factor of 300), which is consistent with the log Bayes factor of 153 obtained by Ellis et al. (2012a). The difference may be caused by the artificially chosen cut-offs of the prior distributions of the strain amplitude and power index, which may make the Bayes factor uncertain (see Section 1.3.2.1).

The posterior probability distributions of the three hyperparameters obtained by applying the Bayesian nonparametric method to the IPTA mock data challenge are presented in Fig. 4.7.

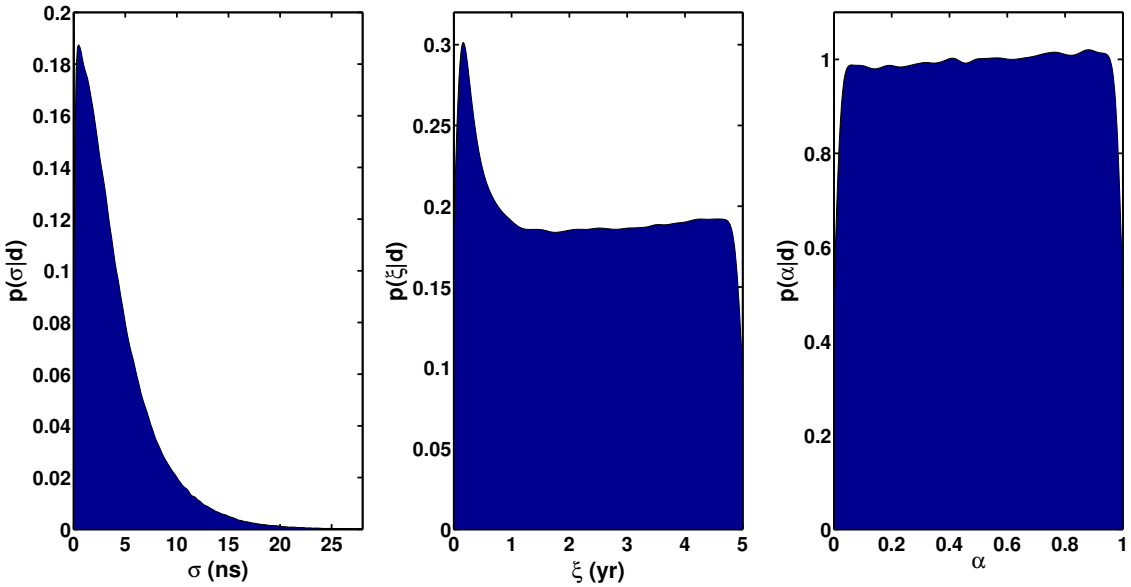


Fig. 4.6 Posterior probability densities of 3 hyperparameters — σ, ξ and α, for analysis on the noise alone data described in Section 4.4.3.3. Their fractional uncertainties are respectively 92.5%, 63.5%, 54.6%.

The previous work (e.g., Ellis et al., 2012a; van Haasteren et al., 2013) characterizes the background by its gravitational wave strain amplitude and the spectrum index. However, the Bayesian nonparametric method characterizes the background by the rms uncertainty of its pulsar timing signal σ, the characteristic timescale of the signal ξ and the shape parameter α. The parameters in the previous work and the hyperparameters in the Bayesian nonparametric method characterize different properties of the background. They may be related but we may not be able to characterize their relationship. Correspondingly, it is difficult to compare the parameter estimations of the two kinds of methods.

4.5 Conclusion

In this chapter, we use a Bayesian nonparametric method to analyze the pulsar timing array data set which may contain a contribution from a gravitational wave background originating

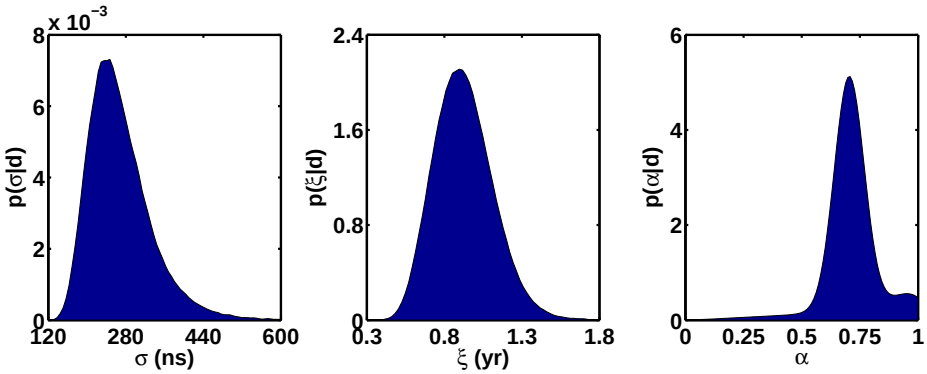


Fig. 4.7 Posterior proability distributions of the three hyperparameters obtained by applying the Bayesian nonparametric method to the closed data set 2 of the IPTA mock data challenge.

from the superposition of gravitational waves generated by supermassive black hole binaries in the universe.

When the number of the gravitational wave sources that significantly contribute to pulsar timing signals is small, the previous methods based on the assumption that the background spectrum is a power law may be restrictive. In order to detect a generic gravitational wave background, we treat it as a deterministic process rather than a stochastic one as before, since each gravitational wave from a single source is a deterministic process. Instead of parameterizing gravitational wave from each single source, we use a different method — Bayesian nonparametrics — to avoid the over-parameterization. In this way, we set strong constraints on the feasible shapes of the pulsar timing residuals induced by the background. We have found that our method works efficiently for theoretically expected signals. When the number of gravitational wave sources becomes small and the assumption of a power law spectrum becomes ineffective, our method is still able to detect and characterize the signal while the conventional method becomes less effective.

For the purpose of demonstration, we apply our Bayesian nonparametric analysis to the pulsar timing data of the 4 best millisecond pulsars in current IPTA, as the capability of detection and characterization of gravitational waves will be dominated by these pulsars (Burt et al., 2011). However, our analysis can be straightforwardly applied to analyze the data of all the pulsars in

the IPTA. In the future, the effective number of pulsars whose timing noises are low enough to detect gravitational waves is expected to significantly increase with the commissioning of more sensitive radio telescopes such as the Five-hundred-meter Aperture Spherical Telescope (Nan et al., 2011) and the Square Kilometer Array (Dewdney et al., 2009). Applying our analysis method to the pulsar timing data collected by these future telescopes will significantly improve the detection sensitivity and interpretation of the signals.

While the context of our discussion focuses on pulsar timing arrays, the analysis itself is directly applicable to detect and characterize any signals that arise from the superposition of a large number of astrophysical events, such as detecting the high frequency gravitational wave background by LIGO (Abadie et al., 2012).

Chapter 5

Searching for Periodic Gravitational Waves with Pulsar Timing Arrays

5.1 Introduction

Gravitational waves in the $\sim$ 3– 30 nHz band generated from supermassive black hole binaries are widely anticipated to produce periodic signals that pulsar timing arrays will investigate. Most previous analyses aimed at detecting such gravitational waves using pulsar timing have argued that, because the contribution to the response from the different pulsar terms is uncorrelated, the "mean" contribution vanishes and the pulsar term contribution to the PTA's overall response can be ignored (Sesana et al., 2009; Yardley et al., 2010; Babak & Sesana, 2012; Ellis et al., 2012b,c). Recent work (e.g., Lee et al., 2011; Ellis, 2013; Arzoumanian et al., 2014) has incorporated the pulsar term into an analysis by treating the pulsar distances as the parameters to be determined. If we can measure the pulsar distances by other techniques such as parallax, such information may help improve the gravitational wave detection. Here we show that the pulsar term contribution to the response need not and should not be ignored: i.e., that, even though it is unknown, it can be included in a Bayesian analysis of PTA data. Doing so leads to a more sensitive analysis and a more accurate estimation of the source parameters than if the contribution is ignored. Furthermore, we will show that we should treat the pulsar term phases but not pulsar distances as the unknown parameters to be determined or marginalized. Pulsar distance measurements will influence gravitational wave analysis only if we can obtain a parsec level precision, which is very difficult to achieve even with our next generation radio telescopes such as the Square Kilometer Array.

In Sections 5.2 and 5.3, we investigate when gravitational wave signals can be treated as periodic in the pulsar timing array waveband. In Section 5.4 we describe our Bayesian analysis methodology. We also show that we should treat the pulsar term phase but not the pulsar distances as the parameters to be determined or marginalized, unless we can achieve parsec distance measurements. In Section 5.5 we illustrate the effectiveness of this analysis by applying it to several representative examples. Finally, we summarize our conclusion in Section 5.6.

5.2 Pulsar timing response to gravitational waves

In this section we review in detail the response of an Earth-pulsar system to gravitational waves generated from circular orbit binaries. Previous calculations of the pulsar timing response to such waves have made numerous approximations or otherwise neglected terms without justification; here we attempt to be more complete. We apply the response to the radiation from evolving circular orbit binary systems, identifying when these may be treated as strictly periodic. Finally, we show how the Earth-pulsar distance that appears in the response may, for the response to plane gravitational waves, be replaced by a phase.

5.2.1 Earth term, pulsar term and pulsar response

Eq. (1.29) gives the gravitational wave induced timing residual $\tau_{\rm GW}$ for the pulse arriving at Earth at time t. The $\mathcal{H}_{(A)}$ in Eq. (1.29) may be written as the sum of three integrals:

$$\mathcal{H}_{(A)}(t) = \int_0^L h_{(A)}(t-(L-\lambda)(1+\hat{k}\cdot\hat{n}))d\lambda \tag{5.1a}$$

$$= -\int_{-L}^{-L+t} \frac{h_{(A)}(\lambda'-L\hat{k}\cdot\hat{n})}{1+\hat{k}\cdot\hat{n}}d\lambda' + \int_{-L}^{L\hat{k}\cdot\hat{n}} \frac{h_{(A)}(\lambda'-L\hat{k}\cdot\hat{n})}{1+\hat{k}\cdot\hat{n}}d\lambda' + \int_0^t \frac{h_{(A)}(\lambda')}{1+\hat{k}\cdot\hat{n}}d\lambda' \tag{5.1b}$$

Even though it is only the sum of these three integrals that has a physical meaning, it has been conventional to give each of these three integrals a specific interpretation:

- The first integral's integrand may be regarded as the wave strain at the location of the pulsar and at the time of pulse emission (pulsar term).
- The second integral is independent of time (though it does depend on the gravitational wave propagation direction, the direction and distance to the pulsar from Earth). Since it does not contribute to the time dependent response it is generally ignored.
- The third integral's integrand may be regarded as the wave strain at Earth (Earth term).

Though it is conventional, the separation of the pulsar timing response to gravitational waves into an "Earth term" and a "pulsar term" contribution is misleading. Only the sum of these three terms — i.e., the total response — is physical.

5.3 Gravitational waves from supermassive black-hole binary systems

5.3.1 Introduction

There is general agreement on the scenario by which the supermassive black hole binaries (Sesana et al., 2008, 2009; Haiman et al., 2009) that PTA observations hope to observe are formed. In this scenario galaxies are built up by mergers between lower mass progenitors, each of which hosts a supermassive black hole progenitor. During the major merger of two such progenitors, dynamical friction drives the two supermassive black holes toward the merged galaxy's nucleus, where they form a bound pair. Dynamical friction, gravitational "slingshot" ejection of stars whose orbits intersect the binary, and gas accretion continue to tighten the binary until the mass enclosed in the orbit is a small fraction of the binary's mass. Gas accretion and the slingshot ejection of stars on centrophilic orbits continue to tighten the binary until gravitational wave radiation reaction

become the dominant dissipation mechanism, at which point merger becomes inevitable (Begelman et al., 1980).

The transition to a gravitational wave driven orbit is expected to take place when the binary separation is in the range 10^2–$10^3\,\mathrm{R_S}$, where $\mathrm{R_S} = 2\mathrm{GM}/c^2$ is the binary's Schwarzschild radius (Milosavljević & Merritt, 2003; Armitage & Natarajan, 2005). For supermassive black holes with masses in the range 10^8 – $10^{11}\,\mathrm{M_\odot}$ this corresponds to orbital periods ranging from months to centuries: i.e., the binary systems that PTA observations seek to observe to may be evolving owing to the ejection of stars on centrophilic orbits, gas accretion, gravitational wave radiation reaction, or some combination of these three mechanisms. Additionally, while orbits driven by gravitational wave radiation reaction rapidly circularize (Peters & Mathews, 1963), gas accretion or gravitational slingshot recoil drive the orbits to high ($e \gtrsim 0.6$–0.8) eccentricity (Sesana, 2010; Preto et al., 2011; Roedig et al., 2011; Khan et al., 2011, 2012b,a).

The binaries that we search for with PTA observations may thus be circular or highly eccentric, evolving at a rate determined by gravitational wave radiation reaction or much more rapidly. Absent a consensus it is necessary to consider all cases; here we focus on the simplest case — circular binary systems driven by gravitational wave radiation reaction — leaving eccentric and evolving binary systems for later study.

5.3.2 Radiation from circular-orbit binary systems

We focus our attention on a supermassive black hole binary with total mass M and symmetric mass ratio $\eta = \mathrm{m_1 m_2}/\mathrm{M}^2$ for component masses $\mathrm{m_1}$ and $\mathrm{m_2}$. In TT-gauge and sufficiently far from the source that the waves may be considered planar over the pulsar-Earth baseline (Deng & Finn, 2011) the quadrupole approximation contribution to the wave propagating in the $\hat{k}$ direction takes

the form

$$h_{+}(u) = \Re\left[e^{2i\psi}\mathfrak{h}(u)\right] \tag{5.2a}$$

$$h_{\times}(u) = \Im\left[e^{2i\psi}\mathfrak{h}(u)\right] \tag{5.2b}$$

where

$$\mathfrak{h}(u) = \mathfrak{h}_{+}(u) + i\mathfrak{h}_{\times}(u) \tag{5.2c}$$

$$\mathfrak{h}_{+}(u) = \frac{1}{2}\frac{\mathcal{M}}{d_L}\left(1+\cos^2\iota\right)\left[\pi f(u)\mathcal{M}\right]^{2/3}\cos(\Phi(u)-\phi_0) \tag{5.2d}$$

$$\mathfrak{h}_{\times}(u) = \frac{\mathcal{M}}{d_L}\left[\pi f(u)\mathcal{M}\right]^{2/3}\cos\iota\,\sin(\Phi(u)-\phi_0) \tag{5.2e}$$

$$f = \frac{1}{2\pi}\frac{d\Phi}{du}, \tag{5.2f}$$

$$\mathcal{M} = M\eta^{3/5} \tag{5.2g}$$

$$u = t - \hat{k}\cdot\vec{x}, \tag{5.2h}$$

$\mathcal{M}$ is the chirp mass ($\mathcal{M} = m_1^{5/3}m_2^{5/3}/\mathrm{M}^{1/5}$), d_L is the source luminosity distance, $\Phi(u)$ is the strain phase, $\cos\iota$ describes the inclination between the orbital angular momentum and radiation propagation direction $\hat{k}$, and ψ describes the orientation of the polarization plane coordinate system.

The phase Φ is determined by the evolution of the binary system, which depends in turn on the mechanism driving the evolution. While uncertainty exists over the dominant evolutionary driver for binaries radiating in the $>$ 1 nHz waveband relevant for pulsar timing observations, gravitational wave radiation reaction sets the floor for the evolution rate. Thus, while we cannot say for certain when binary evolution is *unimportant*, we can say when binary evolution (driven by gravitational wave radiation reaction) is important and cannot be ignored in the evaluation of the

pulsar timing response. For quasi-circular orbits the evolution of the gravitational wave angular frequency $d\Phi/dt$ in the pulsar timing band is dominated by the quadrupole radiation contribution to the evolution:

$$\Phi(u) = -\left[16\eta\left(\pi f M\right)^{5/3}\right]^{-1}. \tag{5.3}$$

Combining Eq. (5.1) and Eq. (5.2), we note that gravitational wave radiation reaction will be important when

$$2\pi \lesssim \frac{\Phi(u)|_{t-L}^{t}}{L\, d\Phi(t)/dt} = \frac{48}{5}\pi^{8/3}\eta L^{2} M^{5/3} f^{11/3} \tag{5.4a}$$

or

$$1 \lesssim 0.3\left(\frac{\eta}{1/10}\right)\left(\frac{M}{10^{9}\,\mathrm{M}_{\odot}}\right)^{5/3}\left(\frac{L}{\mathrm{kpc}}\right)^{2}\left(\frac{f}{3\,\mathrm{yr}^{-1}}\right)^{11/3}. \tag{5.4b}$$

Henceforth we confine our attention to binary systems for which binary evolution is unimportant, meaning at least that inequality Eq. (5.4) is *not* satisfied.

5.3.3 Pulsar timing response to radiation from circular orbit binary systems

Following Eq. (5.2), for circular orbit binary systems whose evolution, over the several thousand year pulsar-Earth light travel time, may be ignored, we write

$$\mathcal{H}_{(+)}(t) = \mathcal{J}_{(+)}(t) - \mathcal{J}_{(+)}(t - L(1 + \hat{k}\cdot\hat{n})) \tag{5.5a}$$

$$\mathcal{H}_{(\times)}(t) = \mathcal{J}_{(\times)}(t) - \mathcal{J}_{(\times)}(t - L(1 + \hat{k}\cdot\hat{n})) \tag{5.5b}$$

with

$$\mathcal{J}_{(+)}(u) = \frac{hI_{+}}{1+\hat{k}\cdot\hat{n}}\cos\left(2\pi f u - (\phi_0 - \phi_{+})\right) \tag{5.5c}$$

$$\mathcal{J}_{(\times)}(u) = \frac{hI_{\times}}{1+\hat{k}\cdot\hat{n}}\sin\left(2\pi f u - (\phi_0 - \phi_{\times})\right) \tag{5.5d}$$

$$h = \frac{h_0}{2}\sqrt{\left(1+\cos^2\iota\right)^2 + 4\cos^2\iota} \tag{5.5e}$$

$$\tan\phi_h = \frac{2\cos\iota}{1+\cos^2\iota} \tag{5.5f}$$

$$I_{+}^2 = \cos^2\phi_h\cos^2 2\psi + \sin^2\phi_h\sin^2 2\psi \tag{5.5g}$$

$$I_{\times}^2 = \sin^2\phi_h\cos^2 2\psi + \cos^2\phi_h\sin^2 2\psi \tag{5.5h}$$

$$\tan\phi_{+} = \tan\phi_h\,\tan 2\psi \tag{5.5i}$$

$$\tan\phi_{\times} = \frac{\tan 2\psi}{\tan\phi_h}. \tag{5.5j}$$

Referring to Eq. (5.5a)–Eq. (5.5d) the pulsar distance contributes to the response through a phase

$$\phi_L = 2\pi f L(1+\hat{k}\cdot\hat{n}) \tag{5.6}$$

that may be reduced modulo 2π. Except when $2\pi f L(1+\hat{k}\cdot\hat{n}) \ll 1$ or the particular pulsar distance is known to much better than the gravitational radiation wavelength we are generally completely ignorant of the value of this phase and may regard it as yet another unknown parameter in our analysis.

5.4 Prior probability distribution

5.4.1 Prior distribution for gravitational wave parameters

Gravitational wave parameters are those characterizing the gravitational waves and are the same for different pulsars. These include 5 parameters — gravitational wave frequency f, amplitude h_0 and its initial phase ϕ_0; its source sky location $\widehat{k}$ and the source orbital orientation $\widehat{\delta}$.

Expectations regarding the distribution of f, h_0 and ϕ_0 are difficult to quantify. We may invoke a model for the source distribution to determine their prior distribution, such as suggested by the theoretical studies of Sesana et al. (2008, 2009). However we proceed, it is helpful to keep in mind that when the observations are conclusive the data dependent term $\Lambda(\mathbf{d}|\boldsymbol{\theta})$ is exponentially sharply peaked. Consequently, as long as the prior probabilities $q_h(h_0)$, $q_f(f)$ and $q_{\phi_0}(\phi_0)$ viewed by themselves do not reflect an overly strong set of expectations it will not matter what particular form they take except at the margins where the observations are suggestive but not conclusive. For the demonstration analysis presented here we choose

$$q_f(f) \sim f^{-1}, \tag{5.7}$$

to represents how the number of sources declines with the increase of the source frequency, and set $q_h(h_0)$ as a uniform distribution between 0 and $+\infty$.

For $q_{\phi_0}(\phi_0)$, we choose a uniform distribution between 0 and 2π, since we do not have any information on the initial phase of the binary system prior to analyzing our observations.

We expect sources to be uniformly distributed across the sky and the source orbital orientation to be uniformly distributed about the circle, so the prior distributions for $\widehat{k}$ and $\widehat{\delta}$ would

be

$$q_{\hat{k}} = q_{\hat{\delta}} = (4\pi)^{-1}. \tag{5.8}$$

5.4.2 Prior distribution for the pulsar term phases

Now we need to determine the prior distribution on the pulsar term phases $\phi_j = 2\pi f L_j(1 + \hat{k}\cdot\hat{n}_j)$. As we discussed, we could treat this phase as an unknown parameter to be determined, and we would marginalize it to obtain the estimation for other gravitational wave parameters. At a first glance, we might expect that the prior distribution on this phase should depend on our knowledge on the pulsar distance L_j. However, we show here that without a parsec level pulsar distance measurement, $2\pi f\sigma_{L_j}(1+\hat{k}\cdot\hat{n}_j) \gg 1$, where σ_{L_j} is the uncertainty of L_j, the prior on the pulsar term phase ϕ_j will be uniform between 0 and 2π.

With a distance measurement uncertainty σ_{L_j}, the prior for L_j is a Gaussian distribution

$$q_{L_j} = \frac{1}{\sqrt{2\pi}\sigma_{L_j}} \exp\left[-\frac{(L_j - L_{0j})^2}{2\sigma_{L_j}^2}\right] \tag{5.9}$$

where L_{0j} is the expection value of the jth pulsar distance. Correspondingly, the prior distribution on the pulsar term phase ϕ_j can be written as

$$q_{\phi_j} = \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left[-\frac{(\phi_j - \phi_{0j})^2}{2\sigma_j^2}\right] \tag{5.10a}$$

where ϕ_{0j} and σ_j respectively denote

$$\phi_{0j} = 2\pi f L_{0j}(1 + \hat{k} \cdot \hat{n}_j) \tag{5.10b}$$

$$\sigma_j = 2\pi f \sigma_{L_j}(1 + \hat{k} \cdot \hat{n}_j). \tag{5.10c}$$

However, if we would like to estimate other gravitational wave parameters, we need to marginalize the posterior probability density over all of these pulsar term phases. The likelihood function depends on the pulsar term phases modulo 2π. Correspondingly, the marginalization of the posterior probability density over the pulsar term phases can be written as

$$\begin{aligned}
&\int P_\theta(\boldsymbol{\theta}|\mathbf{d})\mathrm{d}^j\phi_j \sim \prod_j^{N_p} \int \Lambda_j(\mathbf{d}|\phi_j, \boldsymbol{\theta}) q_{\boldsymbol{\theta}_{\mathrm{GW}}} q_{\phi_j} \mathrm{d}\phi_j \\
&\sim \prod_j^{N_p} \sum_{m_j=-\infty}^{+\infty} \int_0^{2\pi} \Lambda_j(\mathbf{d}|\varphi_j, \boldsymbol{\theta}_{\mathrm{GW}}) q_{\theta_{\mathrm{GW}}} \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left[-\frac{(\varphi_j + 2m\pi - \phi_{0j})^2}{2\sigma_j^2}\right] \mathrm{d}\varphi_j \\
&\sim \prod_j^{N_p} \int_0^{2\pi} \Lambda_j(\mathbf{d}|\varphi_j, \boldsymbol{\theta}_{\mathrm{GW}}) q_{\theta_{\mathrm{GW}}} \vartheta_3\left(\frac{\varphi_j - \phi_{0j}}{2}, \exp(-\sigma_j^2/2)\right) \mathrm{d}\varphi_j
\end{aligned} \tag{5.11a}$$

where $\boldsymbol{\theta}_{\mathrm{GW}}$ denotes the gravitational wave parameters described in the last subsection, φ_j denotes

$$\varphi_j = \phi_j - 2m_j\pi, \tag{5.11b}$$

and m_j is an integer which reduces the pulsar term phase ϕ_j within 0 to 2π. We use the property of the parameter likelihood function

$$\Lambda_j(\mathbf{d}|\phi_j, \boldsymbol{\theta}_{\mathrm{GW}}) = \Lambda_j(\mathbf{d}|\phi_j - 2m_j\pi, \boldsymbol{\theta}_{\mathrm{GW}}) \tag{5.11c}$$

and ϑ_3 is the Jacobi Theta function

$$\vartheta_3(u,q) = 1 + 2\sum_{m=1}^{+\infty} q^{n^2}\cos(2mu). \tag{5.11d}$$

Therefore, the prior distribution on the pulsar term phases ϕ_j is reduced to the prior on φ_j

$$q_\varphi \sim \vartheta_3\left(\frac{\varphi_j - \phi_{0j}}{2}, \exp(-\sigma_j^2/2)\right). \tag{5.12}$$

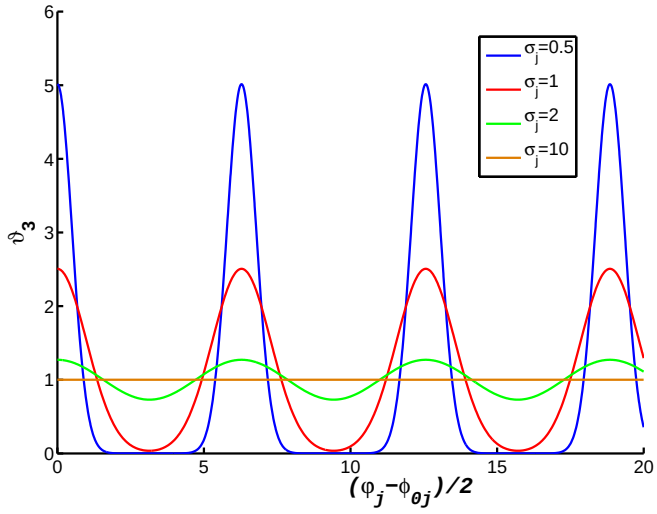


Fig. 5.1 Jacobi Theta function that characterizes the prior distribution of the pulsar term phase, i.e., Eq. (5.12). The prior changes with the different σ_j. It is an approximately uniformly distribution unless $\sigma_j \lesssim 1$.

Figure 5.1 plots the Jacobi Theta function with different σ_j, which characterizes the prior distributions of the pulsar term phase. We can see that unless $\sigma_j \lesssim 1$, this function is approximately a constant. This means unless we can measure the pulsar distance within

$$\sigma_{L_j} = \frac{\sigma_j}{2\pi f(1+\hat{k}\cdot\hat{n})} \lesssim \frac{1}{2\pi f(1+\hat{k}\cdot\hat{n})} \lesssim \text{lyr} \tag{5.13}$$

where the last inequality results from the fact that the source frequency $f \gtrsim 0.1\mathrm{yr}^{-1}$ in the PTA waveband, our knowledge of a pulsar's distance does not affect our knowledge of the pulsar term's phase parameter. For our examples we thus take $q_{\phi_j} = (2\pi)^{-1}$ for each pulsar.

Previous studies state that we can improve the measurement of gravitational wave parameters if we have an independent measurement of pulsar distances and the precision of the pulsar distances do not need to be at the parsec level (Lee et al., 2011; Ellis, 2013). At first glance, this statement sounds correct because additional information of the pulsar distances will help constrain the phases in pulsar terms, which will correspondingly improve the detection sensitivity and estimations of other parameters. However, as we have shown above, without a parsec level pulsar distance measurement, we cannot have an informative prior of the pulsar term phases. This conclusion corrects those previous statements.

5.4.3 Pulsar term is important

Most previous analyses have ignored the pulsar term. The argument for ignoring the pulsar term begins with the observation that the uncertainty in $2\pi f L(1 + \hat{k} \cdot \hat{n})$ is usually much greater than 2π, because in the pulsar timing array waveband, the gravitational wavelength ($\sim$ pc) is much less than the pulsar distance L ($\sim$ kpc). Correspondingly, averaging over all the pulsar terms in a pulsar timing array would make their contribution effectively zero. Thus, the argument goes, the pulsar term can be ignored in the data analysis.

This argument is flawed. We need not and should never average over all the pulsar terms. To search for the evidence of a gravitational wave, we fit a gravitational wave model to all pulsar timing array data but never average the data or the model. The pulsar term contains important information. From Eq. (5.5), we can see that ignoring the pulsar term will lead to a different amplitude and a different phase of the pulsar timing response function. Correspondingly ignoring the

pulsar term would lead to an incorrect timing response function used for matching the gravitational wave signals. The unknown pulsar term phases can be treated as independent parameters to be marginalized or determined.

5.5 Examples

To demonstrate, we apply the analysis method describe above to simulated observations of a monochromatic gravitational wave in the pulsar timing array used in (Demorest et al., 2012). We carry-out two analyses for each data set and compare their results. The first analysis includes marginalizing the pulsar term phases. The other includes only the Earth terms.

5.5.1 Data simulation

We analyze the simulated data containing gravitational wave contributions from three kinds of sources:

(1) A strong source with gravitational wave amplitude signal-to-noise ratio of 20.

(2) A weak source with gravitational wave signal-to-noise ratio of 7.

(3) No signal at all.

For these 3 cases, we will compare the results of including and ignoring pulsar terms. All of the sources are located in the center of the Virgo Cluster, and their gravitational wave frequencies are all 10 nHz. The pulsar timing array we use is listed in Table 3.1. The timing noise we use is described in Section 3.4.2.2.

We uniformly sample 50 observation times across a 5-year duration. The data sets we use are constructed by:

(1) Evaluating the pulsar timing response of each pulsar to the passage of gravitational waves described in the above 3 cases.

(2) Adding the pulsar timing noise to the timing residuals obtained by the first step.

(3) Removing the best-fit linear trend from the noisy timing residuals obtained by the second step because in reality, the linear trend in the data cannot be distinguished from the systematic effects caused by pulsar spin and spin down (Hobbs et al., 2006; Edwards et al., 2006). This step will significantly reduce the noise with long timescale autocorrelation. It will also significantly reduce the likelihood that the timing response model may fit a long period signal.

5.5.2 Results

In the following three subsections, we present the analysis of the three cases above. We investigate whether the gravitational waves are present in the data sets and infer the gravitational wave parameters.

Table 5.1 DIC differences of the method including pulsar term and the one ignoring the pulsar term. "PT" and "no-PT" respectively denote the model including pulsar terms and that ignoring pulsar terms.

Source	SNR	ΔDIC (PT)	ΔDIC (no-PT)	max Z-score (PT)	max Z-score (no-PT)
strong	20	-123	-20	0.289	0.249
weak	7	-11	-3	0.222	0.254
None	0	7	5	0.288	0.289

5.5.2.1 A strong signal

The results of the analysis on gravitational waves from the strong source described above is presented in Fig. 5.2, Fig. 5.3 and the first row in Table 5.1. The evidence for the presence of the signal is much greater when we include the pulsar term phase in the analysis than when it is

ignored. Referring to Fig. 5.2 it is clear that including the pulsar term phase in the analysis leads to significantly better estimation of the gravitational wave parameters: the modes of the distributions for h_0 and f are more accurate for the analysis that includes the pulsar terms than for the analysis that ignores them; and the parameter estimation precision, as measured by the fractional standard deviation in the estimated h_0 and f is 40% better for the analysis that includes the pulsar terms. Referring to Fig. 5.3 we can see that neither of the methods can localize the source within 100deg^2, although the analysis including the pulsar terms offers a better precision than that ignoring the pulsar terms.

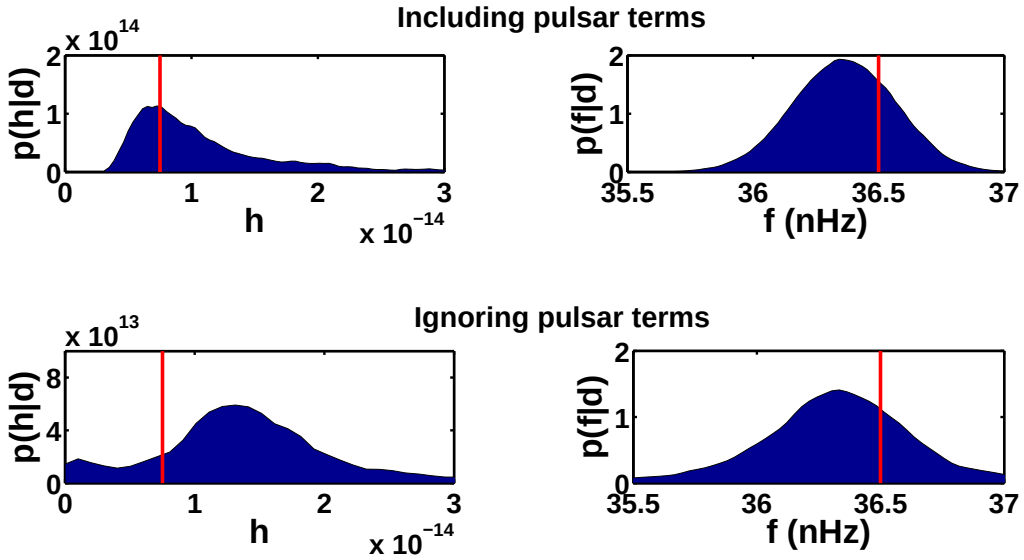
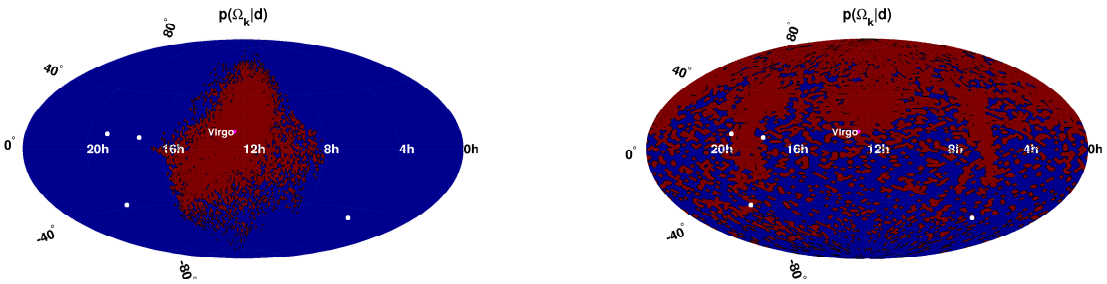


Fig. 5.2 Posterior probability distribution of gravitational wave strain amplitude h_0 and frequency f for the strong source. When including pulsar terms in the analysis, the fractional uncertainty of h_0 and that of f are respectively 20% and 0.5%. However when ignoring pulsar terms, the counterparts of the two fractional uncertainties respectively increase to 35% and 0.8%. The red line denotes the parameters we use to simulate the signal.



(a) Inference of the source sky location by the method including the pulsar terms for the strong source.

(b) Inference of the source sky location by the method ignoring the pulsar terms for the strong source.

Fig. 5.3 Posterior probability distribution of the sky location of the strong source. The inferred sky location has a 99% probability of staying within the red region, and the actual sky location is labelled "Virgo". Including pulsar terms in the analysis localizes the source within $\sim 200\text{deg}^2$. However, ignoring pulsar terms only localizes the source within $\sim 10^4\text{deg}$. The white squares show the locations of the pulsars used as our pulsar timing array.

5.5.2.2 A weak signal

The results of the analysis on gravitational waves from the weak source described above is presented in Fig. 5.4 and the second row in Table 5.1. In this case the contrast between the two analyses is more stark: there is no significant evidence for the presence of a gravitational wave signal, but conclusive evidence when the pulsar term is included in the analysis; and, for the analysis including the pulsar term, the estimated gravitational wave strain amplitude is determined to better than 60% and the signal frequency is determined to better than 3%, but the analysis ignoring the pulsar term cannot accurately determine the gravitational wave parameters.

5.5.2.3 No Signals

The results of the analysis on "no signal" data described above is presented in Fig. 5.5 and the third row in Table 5.1. In this case, neither of the methods shows evidence of gravitational waves.

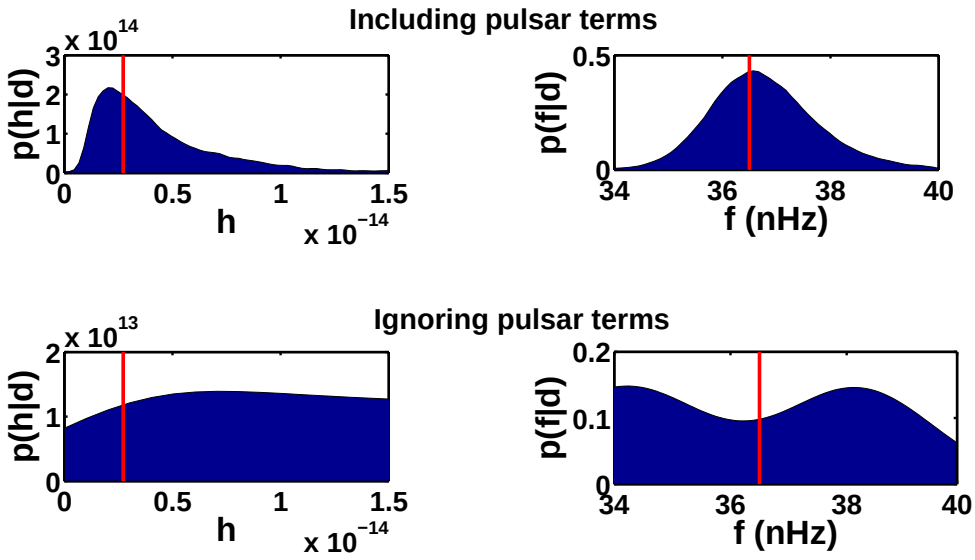


Fig. 5.4 Posterior probability distribution of gravitational wave strain amplitude h_0 and frequency f for the weak source. When including pulsar terms in the analysis, the fractional uncertainty of h_0 and that of f are respectively 60% and 3%. However when ignoring pulsar terms, the counterparts of the two fractional uncertainties both $\gtrsim$ 100%. The red line denotes the parameters we use to simulate the signal.

5.6 Conclusion

In this chapter, we performed a careful Bayesian analysis and showed that pulsar terms should not be ignored in the analysis of searching for periodic gravitational waves in pulsar timing array data. We need not and should never average different pulsar timing response functions over pulsar terms. Instead, we must fit the data collected by all the pulsars in the array together with gravitational wave models, i.e., pulsar timing response functions. Pulsar timing response functions with pulsar terms have different amplitudes and phases than those without pulsar terms. We have found that even for the data containing a weak gravitational wave signal, the method including pulsar terms still offers much better inference than the one ignoring pulsar terms.

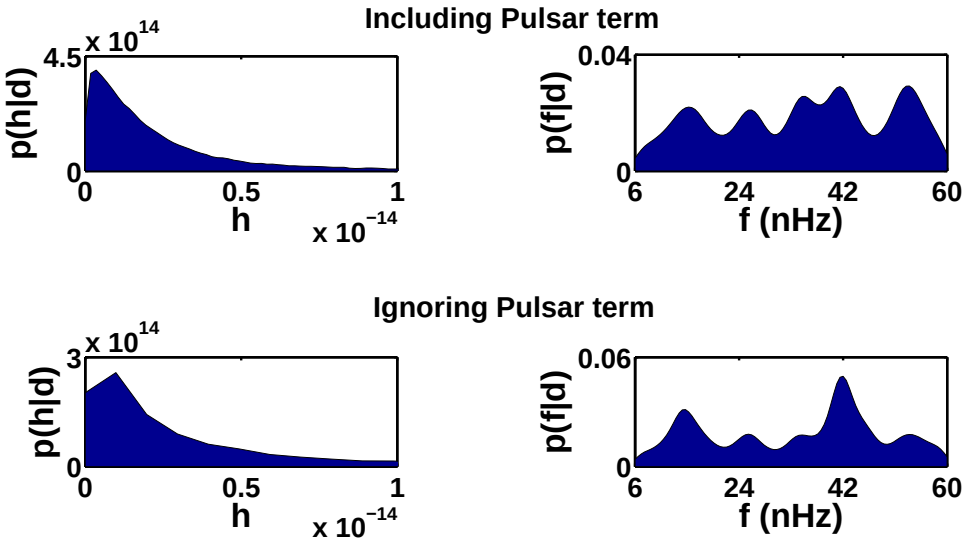


Fig. 5.5 Posterior probability distribution of the gravitational wave strain amplitude h_0 and frequency f.

Furthermore, we have shown that the pulsar distance arises in the response as a phase that, unless the distance is known to parsec scales or better, should be treated as an unknown parameter and marginalized over in the analysis.

Chapter 6

A Planetary System Around the K0 Giant HD 102103

6.1 Introduction

The majority of planet searches to date have focused on FGKM dwarfs (Roques et al., 2014). Given the diversity of planets we have found around these stars, it is important to study planets in a broad range of evolutionary states, to ensure that we have a complete picture of planet formation and evolution. Detecting exoplanets around giant stars via radial velocity observations can increase the population of planets around intermediate mass stars, as the cool atmospheres of the giant stars form narrow spectral lines, making them accessible to radial velocity planet searches, which cannot be achieved by observing the main sequence intermediate mass stars (Gettel, 2012). This chapter describes the analysis of 116 radial velocity (RV) observations of the target HD 102103, a K0 giant star, collected by the Hobby-Eberly telescope (HET). A least squares analysis shows that the data might contain a signal owing to the presence of two exoplanets. However, the existence of the smaller planet is ambiguous because the least squares estimate of its contribution to the RV amplitude is of the same order as the RV residuals[1], which are primarily induced by the p-mode oscillations of the giant star (Lovis & Fischer, 2010). In addition, the least squares estimate of the orbital period of the larger planet turns out to be longer than the observation duration, which makes it difficult to obtain a robust estimate of its orbital eccentricity (Gettel, 2012, Section 2.9). In order to obtain a robust inference from the data, we perform a Bayesian analysis that confirms the presence of two sub-stellar companions and also characterizes the stellar p-mode oscillations. This work is

[1]RV residuals are referred to as the difference between the RV data and the least squares fit of the RV model for all companion objects.

my collaboration with the Penn State-Torun Planet group including S. Gettel, A. Wolszczan, A. Niedzielski, G. Nowak, M. Adamow, P. Zielinski, G. Maciejewski, etc. My contribution is mainly Section 6.4.

6.2 Radial velocity observations of HD 102103

The radial velocity (RV) observations of the target HD 102103 (BD +15 2374, HIP 57320) were made with the Hobby-Eberly Telescope (HET) equipped with the High-Resolution Spectrograph (HRS; Tull, 1998). The observing strategy for the target is to begin with 2 or 3 exposures about 1-2 months apart, and to check for RV variability of greater than 30-50 m s^{-1} and less than $\sim$ 1 km s^{-1}. The RVs were measured at 116 epochs over a period of 2931 days from February 2004 to February 2012. The exposure times ranged between 58 and 240 seconds depending on observing conditions. The signal-to-noise ratio (SNR) values were consistently about 250. The estimated median RV uncertainty for this star was 5 m s^{-1}. The procedure of observing and measurement is described in Gettel et al. (2012a). Fig. 6.2 plots the RVs across the observation period.

6.3 Stellar parameters

The parameters of HD 102013 are summarized in Table 6.1. The detailed procedure of estimating the stellar parameters from spectroscopic data is presented in Gettel (2012, Section 2.9).

6.4 Analysis of radial velocity variations

6.4.1 The difficulty of least squares fit

The RV variations contain the contributions of stellar variability and orbiting stellar or sub-stellar companions. To investigate the stellar and sub-stellar companions, we first remove the

Table 6.1 Stellar Parameters of HD 102103

Parameter	Value
Sp Type	K0 III
$T_{\rm eff}$	4489 ± 25
$\log(L_\star/L_\odot)$	2.13 ± 0.21
$M_\star/M_\odot$	1.7 ± 0.3
$R_\star/R_\odot$	15.3 ± 1.1

Note. "Sp Type" denotes the spectral type of the star; $T_{\rm eff}$ denotes effective temperature; $L_\star$ denotes the stellar luminosity; $M_\star$ denotes the stellar mass; and $R_\star$ denotes the stellar radius.

best-fit linear trend, which corresponds to a long period stellar companion, and then perform a least squares fit of several Keplerian orbits (Gettel, 2012, Section 2.9). It turns out that there are multiple best-fit solutions with period of $\sim$ 490 days, $\sim$ 670 days, $\sim$ 3800 days, etc. (Gettel, 2012, Section 2.9). The rms error of the RV post-fit residuals is $\sim$ 15 m s^{-1}. One or several of the best-fit solutions may fit to the residual noise, which makes the evidence for the existence of the planets ambiguous. This system has been analyzed repeatedly over the course of the observations; however, acquiring further RV measurements has not led to much improvement of the result obtained by the least squares fit (Gettel, 2012, Section 2.9).

6.4.2 Stellar activity

Many stars undergo instrinsic radial velocity variations due to stellar activity. For giant stars, the dominant effects are starspots and non-radial pressure mode (p-mode) oscillations (Gettel, 2012; Niedzielski et al., 2015).

6.4.2.1 Starspots

Stellar rotational modulation due to starspots can mimic radial velocity variations. As a star rotates, a spot on the approaching (blueshifted) side of the star causes less light to be detected from this region, generating a net redshift as we measure the integrated spectrum of the entire

stellar disk. When the spot crosses to the receding (redshifted) side with the stellar rotation, a net blueshift will be generated. Correspondingly, the starspots lead to alternations of the apparent red and blue shifts with the same period as that of the stellar rotation, which confuses the inference of the RV variations. The rotation periods of giant stars can reach several hundreds of days, similar to the typical periods of their planets (Lovis & Fischer, 2010; Gettel, 2012). Therefore, the radial velocity signal induced by a starspot is similar to that of a planet. A representative example of the false detection of a planet is the planet detected around HD 166435 that was later shown to be caused by a starspot (Queloz et al., 2001).

The key difference between the RV variations induced by a starspot and a planet is that unlike the starspot, the planet's orbital motion does not alter the spectral line profiles of the host star. We usually measure the bisector of the line profile at several depths to monitor changes in the line profile. If changes in the line bisector are correlated with radial velocity variations, the signal is likely due to starspots rather than the planet's orbital motion. There are several quasi-periodic models that can be used to fit the starspots variability and minimize the false alarms of planets (Hatzes, 2002; Kipping, 2012).

6.4.2.2 Stellar p-mode oscillations

The stellar p-mode oscillations are excited by turbulent convection near the stellar surface (Christensen-Dalsgaard, 2004). Correspondingly, the excited fluid velocity near the stellar surface will cause a net redshift or blueshift and thus mimic the radial velocity variations (Kjeldsen & Bedding, 1995). For giant stars, the oscillation period of the mode with the maximum power is of the order of hours – days and the mode lifetime is of the order of tens of days (Kjeldsen & Bedding, 1995). The amplitudes of the RV variations induced by the p-mode oscillations are $\sim 10 - 10^2\,\mathrm{m\,s^{-1}}$ (Kjeldsen & Bedding, 1995). The p-mode oscillations are stochastic processes; so

they produce excess stochastic noise in the inferred radial velocity variations (Kjeldsen & Bedding, 1995; Christensen-Dalsgaard, 2004). If we have high precision asteroseismological observations of the stars, we can determine the amplitude and the period of the p-mode oscillations to reduce the false alarms of planets. However, without such information, we are not able to identify which RV variations are due to the p-mode oscillations. Absent this information, previous work has either ignored p-mode oscillations in searching for planetary signals in RV data analysis or just added a simple quadratic term to the RV observations to account for the p-mode oscillations (e.g. Gettel et al., 2012a,b). For planetary RV signals with low amplitudes, however, the p-mode oscillation noise will either hide the signals and raise the false dismissal, or mimic the signals and raise the false alarms. A notable example is the K0 giant HD 102103.

6.4.3 Bayesian analysis

Due to the indefinite conclusion drawn by the least squares fit, we perform a careful Bayesian analysis on the RVs of HD 1020103 in an attempt to determine what contributes to the RVs.

6.4.3.1 Modeling stellar variability

We first model the stellar variability that contributes to the post-fit residuals of the RVs. The dominant contributions to the apparent RV residual for giant stars are stellar p-mode oscillations (Lovis & Fischer, 2010). The radial velocity rms amplitude σ induced by stellar oscillation linearly scales with the luminosity-to-mass ratio (Kjeldsen & Bedding, 1995):

$$\sigma \sim 0.23 \left(\frac{L_\star}{L_\odot}\right) \left(\frac{M_\star}{M_\odot}\right)^{-1} \mathrm{m\,s^{-1}}. \tag{6.1}$$

Based on Table 6.1, the RV rms amplitude of the red giant HD 102103 is $\sim$ 18m s^{-1}, which is consistent with the rms error of the RV residuals. The maximum power of the stellar p-mode

oscillation is at a frequency of (Kjeldsen & Bedding, 1995)

$$f_0 \sim 3 \left(\frac{M_\star}{M_\odot}\right) \left(\frac{R_\star}{R_\odot}\right)^{-2} \left(\frac{T_{\rm eff}}{5777{\rm K}}\right)^{-1/2} {\rm mHz}. \tag{6.2}$$

Correspondingly, according to Table 6.1, the peak oscillation frequency and peak oscillation period of the giant HD 102103 are respectively $\sim 30\mu$Hz and ~ 8hr.

The power in the stellar p-mode oscillation is concentrated at frequencies close to the peak oscillation frequency f_0 (i.e., near resonance) due to the long damping time of the oscillation (Anderson et al., 1990). The power spectral density of the stellar p-mode oscillation near the resonance is approximately modeled as a displaced Cauchy distribution (Anderson et al., 1990), i.e.,

$$S(f) = \frac{\sigma^2}{1 + Q^2(f - f_0)^2/f_0^2} \quad \text{with} \quad f > 0. \tag{6.3}$$

This is derived from the stochastic damped oscillator model (Christensen-Dalsgaard, 2004), where σ is the rms amplitude, f_0 is the peak oscillation frequency and Q is the quality factor that characterizes the number of oscillation cycles before it is significantly damped.

The corresponding covariance matrix for this power spectral density is its cosine transformation:

$${\rm C}_{ij} = \sigma^2 \exp\left(-\frac{2\pi f_0 |t_i - t_j|}{Q}\right) \cos(2\pi f_0 |t_i - t_j|). \tag{6.4}$$

Correspondingly, we model the stellar p-mode oscillation as a Gaussian process with zero mean and covariance matrix Eq. (6.4). The model has three parameters — rms amplitude σ, quality factor Q and peak oscillation frequency f_0.

6.4.3.2 Bayesian inference

We use Bayesian analysis to estimate the p-model oscillation parameters and the Kepler orbital parameters. We choose diffuse prior distributions for all parameters as we do not know much information about the target. All the prior distributions are listed in Table 6.2.

Table 6.2 Prior distributions of the parameters characterizing the stellar p-mode oscillation and the two Kepler orbits. The $\mathcal{U}_{[a,b]}$ denotes the uniform distribution on the interval of lower bound a and upper bound b.

Parameter	Prior distribution
rms p-mode amplitude (m s^{-1})	$\mathcal{U}_{[0,+\infty]}$
quality factor	$\mathcal{U}_{[50,500]}$
peak p-mode period (hr)	$\mathcal{U}_{[0.5,15]}$
orbital semi-amplitudes (m s^{-1})	$\mathcal{U}_{[0,500]}$
orbital periods (yr)	$\mathcal{U}_{[0.1,40]}$
eccentricities	$\mathcal{U}_{[0,1)}$
arguments of periapse (rad)	$\mathcal{U}_{[0,2\pi)}$
phases of periastron passage (rad)	$\mathcal{U}_{[0,2\pi)}$

To investigate the astrophysical contribution to the RV variations, we compare three models of the RV variations and check which one is favored by the data:

1. RV variations are induced by a linear trend, a longer period Keplerian orbit, a shorter period Keplerian orbit, the stellar p-mode oscillation and measurement errors.

2. RV variations are induced by a linear trend, a longer period Keplerian orbit, the stellar p-mode oscillation and measurement errors. The prior distribution of the orbital period in this model is chosen to be $\mathcal{U}_{[5,40]}$ with the unit of years.

3. RV variations are induced by a quadratic trend, a shorter period Keplerian orbit, the stellar p-mode oscillation and measurement errors. The prior distribution of the orbital period in this model is chosen to be $\mathcal{U}_{[0.1,5]}$ with the unit of years.

Model 2 is intended to check for the presence of the shorter period orbit and Model 3 is intended to check for the presence of the longer period orbit. We do not need to check the null hypothesis that the RV variations are induced by a quadratic trend, the stellar p-mode oscillation and measurement errors. This is because the post-fit residuals after removing the best-fit quadratic trend are not a zero mean stationary process, which indicates there are other effects contributing to the RV variations (Gettel, 2012, Section 2.9).

6.4.3.3 Results

In this subsection, we present our results on detecting evidence of exoplanets in the RV data and estimating their parameters.

Table 6.3 summarizes the DICs of the three models listed above, and Fig. 6.1 presents the normalized posterior probability distributions of semi-amplitudes, orbital periods and eccentricities for the two Keplerian orbits, and the rms, quality factor and oscillation period of the stellar p-mode oscillation in the three models.

Table 6.3 DICs of the three models and the maximum absolute value of the Z-score of the convergence diagnostics of the MCMC samples. We have found all the DICs change by less than 4 when we change the prior distributions of the quality factor and the peak oscillation period listed in Table 6.2 respectively to $\mathcal{U}_{[10,800]}$ and $\mathcal{U}_{[0.1,25]}$. This is because the DICs only explicitly depend on the posterior distributions but not the priors. The posterior distributions presented in Fig. 6.1 are significantly different from the uniform priors we assigned, which means it is the likelihood functions but not the priors that primarily determine the outcomes of DICs and parameter estimations.

Model	DIC	max Z-score
1	740	0.208
2	772	0.250
3	760	0.294

We can see that the RV data strongly favors the first model, i.e., the RV variations are induced by a linear trend, a longer period Keplerian orbit, a shorter period Keplerian orbit, the stellar p-mode oscillation and measurement errors. The maximum a posteriori probability estimates

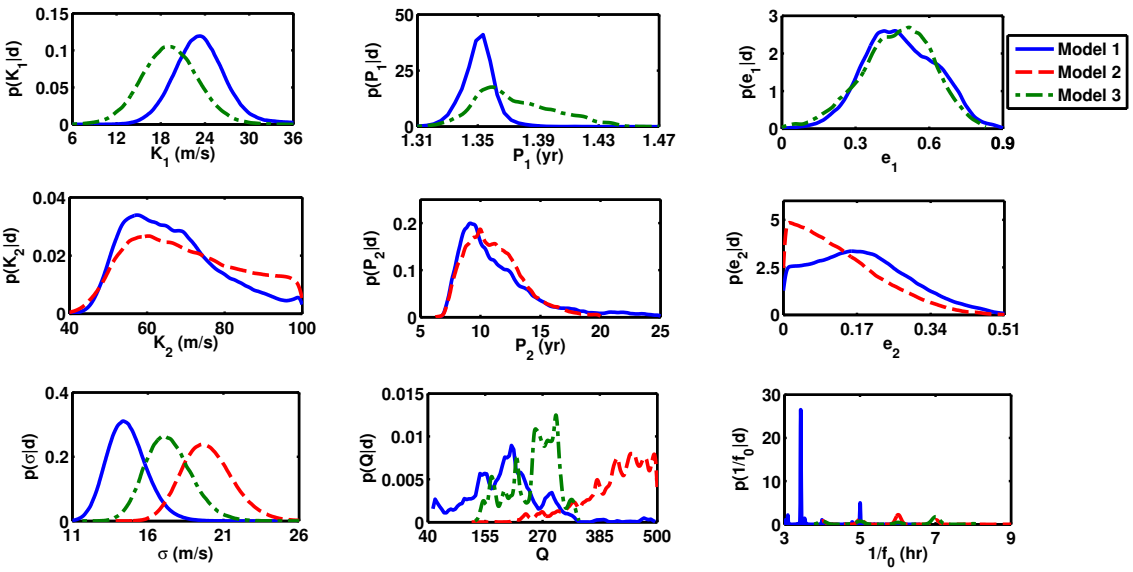


Fig. 6.1 Posterior probability distributions of parameters used in the three models. The K_1, P_1, e_1 and K_2, P_2, e_2 are respectively the semi-amplitude, orbital period and eccentricity of the shorter period orbit and those of the longer period orbit. The σ, Q and $1/f_0$ are respectively rms amplitude, quality factor and peak period of the stellar p-mode oscillation. The solid line refers to model 1, the dash line refers to model 2 and the dash-dot line refers to model 3. At the MAPs, the parameters of the two Keplerian orbital parameters and those of the p-mode oscillation are statistically independent.

(MAPs) and the standard deviations of the Keplerian orbital parameters and the p-mode oscillation parameters are presented in Table 6.4. Best fit models based on MAPs and the RV residuals are presented in Fig. 6.3. The MAPs of the p-mode oscillation parameters are consistent with the order-of-magnitude estimates by using stellar parameters discussed in Section 6.4.3.1. This means that the information of the stellar properties obtained by our Bayesian analysis on the RV observations of the HD 102103 is consistent with that obtained by stellar photometric and spectroscopic observations.

6.4.3.4 Summary

By performing a careful Bayesian analysis, we simultaneously find robust evidence of two planets and characterize the p-mode oscillation of the giant HD 102103. Therefore, it is now very clear that besides a linear trend, there are two sub-stellar companions and a p-mode oscillation

Table 6.4 MAPs and uncertainties of the Keplerian orbital parameters and stellar p-mode oscillation parameters

Orbital Parameters	Short Period Orbit	Long Period Orbit
semi-amplitude (m s^{-1})	23 ± 4	69 ± 12
period (days)	493 ± 4	3833 ± 620
eccentricity	0.46 ± 0.15	0.18 ± 0.11
argument of periastron (deg)	326 ± 25	88 ± 45
initial epoch (MJD)	53200 ± 60	55900 ± 700
minimum mass (M_J)	0.8	5.0
P-mode Parameters		
rms of p-mode (m s^{-1})	15 ± 2	
quality factor	200 ± 100	
peak p-mode period (hr)	3.5 ± 0.4	

that contribute to the RV variations. The linear trend in the RVs is induced by a long-period stellar companion and we can constrain its mass by examining the stellar spectra (Gettel, 2012). The stellar template of the HD 102103 star does not show lines from a companion. Given a S/N of 420, the luminosity ratio of the two objects is of order 8 magnitude and the maximum mass of a M-dwarf companion is $0.5M_{\odot}$ (Gettel, 2012, Section 2.9).

We also test whether this two-planet system is dynamically stable to further check if the planetary system is physical, since planets in an unstable gravitational system should not follow stable Keplerian orbits. We use the Mercury code [2], which is a Fortran program that performs N-body integration of the gravitational system, to test the stability of the system. Prof. Aleksander Wolszczan has run the code and found that this planetary system is stable for $\gtrsim$ 1 million years, which is long enough to sustain the stable Keplerian orbits of the two planets.

6.5 Conclusion

Radial velocity observations of HD 102103 lead to a discovery of two sub-stellar companions and a stellar companion around the host K0 giant star. There are seven (including HD 102103)

[2] url: http://www.arm.ac.uk/∼jec/home.html

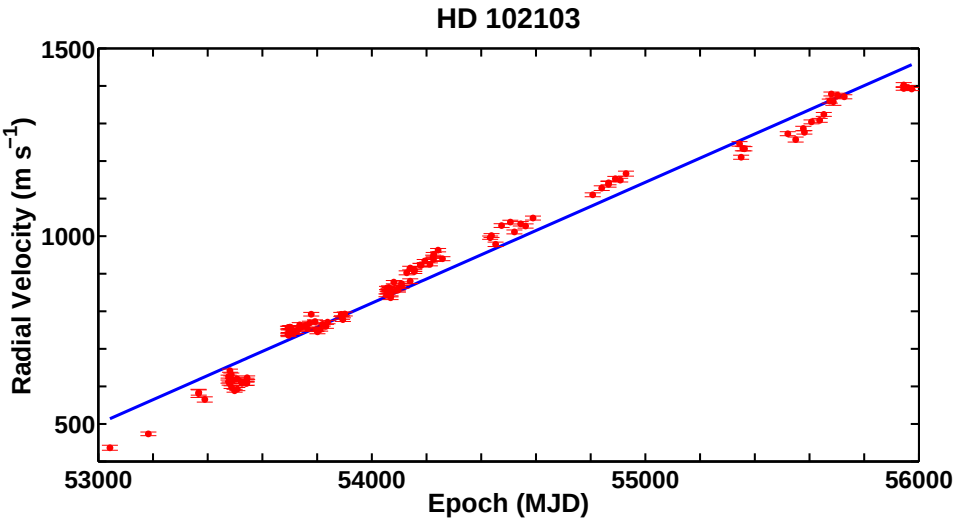


Fig. 6.2 Radial velocity data of HD 102103 and the best fit linear trend that characterizes the RV induced by the Keplerian orbit of the long period stellar companion.

known multiple planetary systems around giants (Niedzielski et al., 2015) and HD 102103 is the only one that contains a stellar companion. The two planets orbiting HD 102103 have orbital periods of 493 and 3833 days and minimum masses of 0.8 M_J and 5.0 M_J, respectively. The planet with the shorter orbital period is one of the least massive planets and also with one of the most eccentric orbits around giant stars (Roques et al., 2014; Niedzielski et al., 2015).

For giant stars, the RV residuals are primarily induced by stellar oscillation (Lovis & Fischer, 2010). However, in the absence of accurate asteroseismology data such as those collected by precision photometry, we do not have a predetermined amplitude or a predetermined period of the stellar oscillation. Correspondingly, it is difficult to confirm the presence of the planets whose signals are masked by the undetermined stellar p-mode oscillation. In order to obtain a robust inference, we model the RV data as the sum of contributions induced by the companion objects and the stellar p-mode oscillation. We then use a Bayesian analysis method to simultaneously infer all the unknown parameters including those characterizing the Keplerian orbits of the companions and those characterizing the stellar p-mode oscillation. Consequently, we confirm the presence of the two planets and accurately estimate their Keplerian parameters. As a bonus, we also estimate

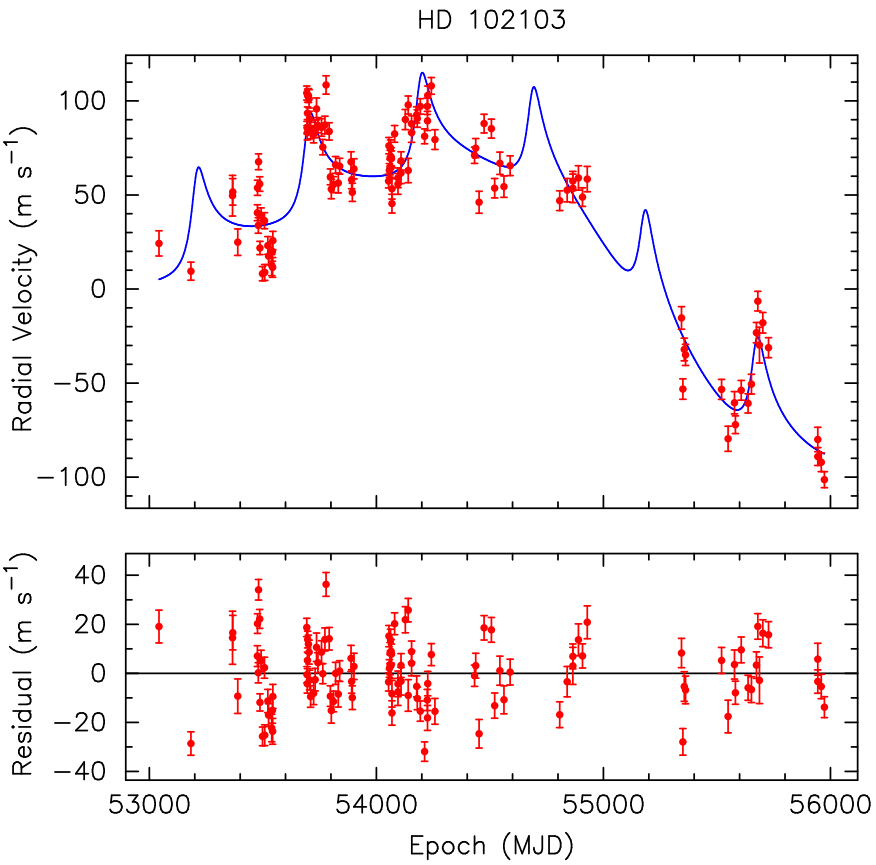


Fig. 6.3 Top: RV measurements and the best-fit two-orbit model after removing the linear trend. Bottom: post-fit residuals.

the amplitude, peak period and quality factor of the stellar oscillation with reasonable precision. Although this information may not be precise enough to study stellar asteroseismology, it can help to check the consistency of the stellar parameters measured by stellar photometry and spectroscopy.

Chapter 7

Conclusions and Prospects

This thesis contains five projects — one on measuring gravitational wave timing parallax via pulsar timing arrays to determine the source luminosity distance; four on developing Bayesian analysis methods to search for gravitational waves in the pulsar timing array data and also search for exoplanets around giant stars in radial velocity observations. I have presented two Bayesian nonparametric methods that can be used to search for gravitational wave bursts and the gravitational wave background from the pulsar timing array data. I have also presented a Bayesian parametric method that can be used to search for gravitational waves generated from a slowly evolving supermassive black hole binary. I have applied the methods of gravitational wave data analysis to search for exoplanets around a K0 giant HD 102013 from radial velocity data collected by the Hobby-Eberly Telescope. Consequently, I have confirmed the presence of two Jupiter mass planets around the giant, which cannot be achieved by the conventional least squares analysis.

7.1 Bayesian analysis on searching for gravitational waves in the pulsar timing array data

Direct detection of gravitational waves will open a new window of our universe complementary to the conventional electromagnetic observations. Looking through this window requires advanced data analysis methodology and techniques. In Chapter 2, I describe a technique that uses the pulsar timing array to measure the phasefront curvature of a gravitational wave generated from a nearby source, which can help to determine the luminosity distance of the source. This is an example of how detecting gravitational waves can be used as a tool to study our universe. From

Chapter 3-5, I describe Bayesian data analysis methods that can be used to search for gravitational waves from different kinds of sources in the pulsar timing array data.

In Chapter 3, I propose a new method of Bayesian nonparametric analysis, aimed at detecting gravitational wave bursts in the pulsar timing array waveband. Instead of fitting the data with some specific physical model, the method tries to set constraints on the features of the pulsar timing response to a gravitational wave burst by a Gaussian process prior and search for the feasible signal patterns in the data. I apply the method to the simulated pulsar timing array data to show that the Bayesian nonparametric analysis can investigate whether a gravitational wave burst is present in the data and also estimate the important source parameters such as burst duration, even though we may not be able to model the source dynamics with some specific physical process.

In Chapter 4, I propose another Bayesian nonparametric method aimed at searching for a generic gravitational wave background in the pulsar timing array band. Unlike previous methods all assuming that the background is a Gaussian distributed stochastic process (e.g., Jenet et al., 2005; van Haasteren et al., 2009), I treat it as a deterministic process and assign a Gaussian process prior to constrain the shape of the pulsar timing signal produced by the background. In this way, my method can search for the only feasible patterns of the signal in the data but not make any assumptions of the distribution of the background. I apply this method to the simulated pulsar timing array data and demonstrate that my method is more effective than the previous methods when the non-Gaussianity of the gravitational wave background is significant.

In Chapter 5, I propose a Bayesian parametric method aimed at detecting gravitational waves from a slowly evolving supermassive black hole binary. The pulsar timing response to such gravitational waves is the difference of two sinusoidal functions with the same frequency. The two functions are referred to as the Earth term and the pulsar term. Most previous methods (e.g., Sesana et al., 2009; Yardley et al., 2010; Babak & Sesana, 2012; Ellis et al., 2012b,c) ignore

the pulsar terms in the analysis by claiming that they will be effectively zero when averaging over all pulsars in the array. Recent work (e.g., Lee et al., 2011; Ellis, 2013; Arzoumanian et al., 2014) has included the pulsar terms in the analysis by treating the pulsar distances to the Earth as the parameters to be determined, because in this way we may improve the inference of the gravitational waves by constraining the pulsar distances with other independent measurements. I show that the pulsar terms are important to the gravitational wave data analysis and including them will improve the detection sensitivity and also the estimation precision of gravitational wave parameters. Furthermore, I demonstrate that we should use the pulsar term phases but not pulsar distances as the unknown parameters to be determined, unless we can measure the distance at the parsec level precision. Without such high precision distance measurements, we have to face the difficulty of computing a multi-mode posterior probability distribution if we use pulsar distances as the unknown parameters but will not improve the inference compared with that of treating the pulsar term phases as parameters to be determined.

All the methods I have developed in this thesis can be used to search for gravitational waves in the IPTA data. In particular, the Bayesian nonparametric method to detect the gravitational wave background drives interest in the IPTA data analysis group. They now start to use it to search for an isotropic gravitational wave background in the current pulsar timing data collected by the Arecibo telescope, the Greenbank telescope and the Parkes telescope.

7.2 Bayesian analysis on searching for exoplanets around giant stars

There is a small but growing number of planet detections around giant stars. Studying this population will complement our knowledge of the planets around FGKM dwarfs and ensure a complete picture of planet formation and evolution. Giant stars experience strong p-mode oscillations that will produce stochastic noise in the radial velocity data. The stellar noise may be so strong

that it can mask the signals induced by the surrounding planetary orbits in the radial velocity data. HD 102103 is one of the giant stars whose radial velocity observations contain the contribution of their p-mode oscillation. This makes it difficult to search for the planets around the giant in the radial velocity data and the least squares analysis is not able to confirm the existence of the planets.

I have developed a Bayesian analysis method that helps to provide a robust inference of the radial velocity observations of HD 102103. I model the stellar p-mode oscillation as a Gaussian stochastic process, and model the data as the sum of the p-mode oscillation and planetary signals. As a consequence, I confirm the existence of two planets — one with a minimum mass of $0.8\,\mathrm{M}_J$, an orbital period of 493 days and an orbital eccentricity of 0.46; the other one with a minimum mass of $5.0\,\mathrm{M}_J$, an orbital period of 3833 days and an orbital eccentricity of 0.18. There is also a M dwarf companion in the system with a maximum mass of $0.5\,\mathrm{M}_\odot$ and an orbital period much longer than the observation duration. This is the first discovered planetary system around a giant star with a stellar companion.

As a bonus, my method also provides a reasonable inference of the parameters characterizing the stellar p-mode oscillation, i.e., the rms amplitude, the quality factor and the peak period of the oscillation. Although the estimation is not so precise as other techniques such as high precision photometry, in the case of a lack of precise asteroseismological observations of HD 102103, it can help check the consistency of the information on the stellar properties obtained by the stellar photometry and spectroscopy.

So far a lot of giant star targets observed by the Penn State-Torun Planet Search group have not been analyzed in detail because the least squares analysis failed to distinguish the planetary signals from the stellar p-mode oscillations. However, the successful confirmation of planets around HD 102103 demonstrates the effectiveness of the Bayesian analysis to provide robust inferences of the radial velocity observations of such targets. Correspondingly, it is worthwhile to apply the

Bayesian analysis method I developed to re-analyze those data in detail, and it is possible that my method may significantly enlarge the sample of planets around giant stars. Therefore, in the future I will develop my Bayesian method into a general pipeline to search for exoplanets around giant stars in the radial velocity data.

Bibliography

Aasi, J. et al. 2013, Phys. Rev. D, 87, 022002

Abadie, J. et al. 2012, Phys. Rev. D, 85, 122001

Abramowitz, M. & Stegun, I. A., eds. 1964, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables (Mineola, NY: Dover)

Adamów, M., Niedzielski, A., Villaver, E., Wolszczan, A., & Nowak, G. 2014, Astron. & Astrophys., 569, A55

Adler, R. J. 1981, The Geometry of Random Fields, Wiley Series in Probability and Mathematical Statistics (Chichester, UK: John Wiley & Sons)

Akaike, H. 1974, IEEE Transactions on Automatic Control, 19, 716

Anderson, E. R., Duvall, T. L., & Jefferies, S. M. 1990, Astrophys. J, 364, 699

Armitage, P. J. & Natarajan, P. 2005, Astrophys. J, 634, 921

Arzoumanian, Z. et al. 2014, Astrophys. J, 794, 141

Babak, S. & Sesana, A. 2012, Phys. Rev. D, 85, 044034

Begelman, M. C., Blanford, R. D., & Rees, M. J. 1980, Nature, 287, 307

Bernardo, J. M. & Smith, A. F. M. 2003, Bayesian Theory, Wiley Series in Probability and Mathematical Statistics (Chichester, UK: John Wiley & Sons)

Bhat, N. D. R., Cordes, J. M., Camilo, F., Nice, D. J., & Lorimer, D. R. 2004, Astrophys. J, 605, 759

Blanchet, L. 2006, Living. Rev. Relativity, 9, 4

Bretthorst, G. L. 1988, Bayesian Spectrum Analysis and Parameter Estimation, Springer Series in Statistics (New York, NY: Springer)

Burt, B. J., Lommen, A. N., & Finn, L. S. 2011, Astrophys. J, 730, 17

Cenko, S. B. et al. 2011, Astrophys. J, 732, 29

Christensen, N. 1992, Phys. Rev. D, 46, 5250

Christensen-Dalsgaard, J. 2004, Solar. Phys, 220, 137

Corbin, V. & Cornish, N. J. 2010, Pulsar Timing Array Observations of Massive Black Hole Binaries, arXiv: 1008.1782

Cordes, J. M. & Jenet, F. A. 2012, Astrophys. J, 752, 54

Damour, T. & Vilenkin, A. 2001, Phys. Rev. D, 64, 064008

Deck, K. M. et al. 2012, Astrophys. J, 755, L21

Deller, A. T. 2009, PhD thesis, Swinburne University of Technology, Melbourne, Victoria

Deller, A. T., Verbiest, J. P. W., Tingay, S. J., & Bailes, M. 2008, Astrophys. J, 685, L67

Demorest, P. B. et al. 2012, Astrophys. J, 762, 94

Deng, X. 2014, Phys. Rev. D, 90, 024020

Deng, X. & Finn, L. S. 2011, Mon. Not. R. Astron. Soc., 414, 50

Detweiler, S. 1979, Astrophys. J, 234, 1100

Dewdney, P. E., Hall, P. J., Schilizzi, R. T., & Lazio, T. J. L. W. 2009, Proc. IEEE, 97, 1482

Doksum, K. A. 1974, The Annals of Statistics, 2, 183

Edwards, R. T., Hobbs, G. B., & Manchester, R. T. 2006, Mon. Not. R. Astron. Soc., 372, 1549

Ellis, J. 2013, Class. Quant. Grav, 30, 224004

Ellis, J., Siemens, X., & Chamberlin, S. 2012a, Results of the First IPTA Closed Mock Data Challenge, arXiv: 1210.5274

Ellis, J. A., Jenet, F. A., & McLaughlin, M. A. 2012b, Astrophys. J, 753, 96

Ellis, J. A., Siemens, X., & Creighton, J. D. E. 2012c, Astrophys. J, 756, 175

Feller, W. 1945, Bull. Amer. Math. Soc., 51, 800

Ferguson, T. S. 1973, The Annals of Statistics, 1, 209

Finn, L. S. 1985, Class. Quant. Grav, 2, 381

—. 2009, Phys. Rev. D, 79, 022002

Finn, L. S. & Lommen, A. N. 2010, Astrophys. J, 718, 1400

Flanagan, E. E. 1993, Phys. Rev. D, 48, 2389

Foster, R. S. & Backer, D. C. 1990, Astrophys. J, 361, 300

Gelman, A. 2006, Bayesian. Anal., 1, 515

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. 2004, Bayesian Data Analysis, 2nd edn., Texts in Statistical Science (Boca Raton, FL: Chapman & Hall/CRC)

Gettel, S. 2012, PhD thesis, The Pennsylvania State University, University Park, PA

Gettel, S., Wolszczan, A., Niedzielski, A., Nowak, G., Adamow, M., Zielinski, P., & Maciejewski, G. 2012a, Astrophys. J, 756, 53

—. 2012b, Astrophys. J, 756, 53

Geweke, J. 1992, in Bayesian statistics, ed. J. Bernardo, J. Berger, A. Dawid, & A. Smith (Oxford, United Kingdom: Oxford University Press)

Ghosh, J. K. & Ramamoorthi, R. V. 2003, Bayesian Nonparametrics, Springer Series in Statistics (New York, NY: Springer-Verlag)

Gibbons, J. D. & Chakraborti, S. 2003, Nonparametric Statistical Inference, 4th edn. (Boca Raton, FL: Chapman & Hall/CRC)

Haiman, Z., Kocsis, B., & Menou, K. 2009, Astrophys. J, 700, 1952

Hastings, W. K. 1970, Biometrika, 57, 97

Hatzes, A. P. 2002, Astronomische Nachrichten, 323, 392

Hellings, R. W. & Downs, G. S. 1983, Astrophys. J, 265, L39

Hjort, N. L., Holmes, C., Müller, P., & Walker, S. G., eds. 2010, Bayesian Nonparametrics, Cambridge Series in Statistical and Probabilistic Mathematics (Cambridge, UK: Cambridge University Press)

Hobbs, G. B., Edwards, R. T., & Manchester, R. T. 2006, Mon. Not. R. Astron. Soc., 369, 655

Hobbs, G. B. et al. 2010, Class. Quant. Grav, 27, 084043

Jeffreys, H. 1946, Proc. R. Soc. London., Ser. A, 186, 453

Jenet, F. A., Armstrong, J. W., & Tinto, M. 2011, Phys. Rev. D, 83, 081301

Jenet, F. A., Hobbs, G. B., Lee, K. J., & Manchester, R. N. 2005, Astrophys. J, 625, L123

Jenet, F. A., Lommen, A. N., Larson, S. L., & Wen, L. 2004, Astrophys. J, 606, 799

Jenet, F. A. et al. 2006, Astrophys. J, 653, 1571

Joanes, D. N. & Gill, C. A. 1998, J. R. Stat. Soc. D., 47, 183

Kass, R. E. & Raftery, A. E. 1995, J. Am. Stat. Assoc., 90, 773

Khan, F. M., Berentzen, I., Berczik, P., Just, A., Mayer, L., Nitadori, K., & Callegari, S. 2012a, Astrophys. J, 756, 30

Khan, F. M., Just, A., & Merritt, D. 2011, Astrophys. J, 732, 89

Khan, F. M., Preto, M., Berczik, P., Berentzen, I., Just, A., & Spurzem, R. 2012b, Astrophys. J, 749, 147

Kipping, D. 2012, Mon. Not. R. Astron. Soc., 427, 2487

Kjeldsen, H. & Bedding, T. R. 1995, Astron. & Astrophys., 293, 87

Kullback, S. & Leibler, R. A. 1951, Ann. Math. Stat., 22, 79

Leblond, L., Shlaer, B., & Siemens, X. 2009, Phys. Rev. D, 64, 064008

Lee, K. J., Bassa, C. G., Janssen, G. H., Karuppusamy, R., Kramer, M., Liu, K., Perrodin, D., Smits, R., & Stappers, B. W. 2014, Mon. Not. R. Astron. Soc., 441, 2831

Lee, K. J., Wex, N., Kramer, M., Stappers, B. W., Bassa, C. G., Janssen, G. H., Karuppusamy, R., & Smits, R. 2011, Mon. Not. R. Astron. Soc., 414, 3251

Lentati, L., Alexander, P., Hobson, M. P., Taylor, S., Gair, J., Balan, S. T., & van Haasteren, R. 2013, Phys. Rev. D, 87, 104021

Lentati, L., Hobson, M. P., & Alexander, P. 2014, Mon. Not. R. Astron. Soc., 444, 3863

Lommen, A. N. 2001, PhD thesis, University of California, Berkeley, Berkeley, CA

Lommen, A. N. & Backer, D. C. 2001, Astrophys. J, 562, 297

Lovis, C. & Fischer, D. 2010, in Exoplanets, ed. S. Seager (Tucson, AZ: the University of Arizona Press)

Lyne, A., Hobbs, G., Kramer, M., Stairs, I., & Stappers, B. 2010, Science, 329, 408

Lyne, A. G. & Graham-Smith, F., eds. 2006, Pulsar Astronomy, 3rd edn. (Cambridge, UK: Cambridge University Press)

Maggiore, M. 2007, Gravitational Waves (Oxford, UK: Oxford University Press)

Manchester, R. N., Hobbs, G. B., Teoh, A., & Hobbs, M. 2005, Astron. J., 125, 1993

Manchester, R. N. et al. 2013, PASA, 30, e017

Meng, X. & Rubin, D. B. 1992, Biometrika, 79, 103

Metroplis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. 1953, J. Chem. Phys., 21, 1087

Milosavljević, M. & Merritt, D. 2003, Astrophys. J, 596, 860

Mingarelli, C. M. F., Sidery, T., Mandel, I., & Vecchio, A. 2013, Phys. Rev. D, 88, 062005

Misner, C. W., Thorne, K. S., & Wheeler, J. A. 1973, Gravitation (New York, NY: W. H. Freeman and Company)

Nan, R., Li, D., Jin, C., Wang, Q., Zhu, L., Zhu, W., Zhang, H., Yue, Y., & Qian, L. 2011, Int. J. Mod. Phys. D, 20, 989

Neyman, J. & Pearson, E. 1933, Philos. Trans. R. Soc. London., Ser. A, 231, 289

Niedzielski, A., Konacki, M., Wolszczan, A., Nowak, G., Maciejewski, G., Gelino, C. R., Shao, M., Shetrone, M., & Ramsey, L. W. 2007, Astrophys. J, 669, 1354

Niedzielski, A., Villaver, E., Wolszczan, A., Adamów, M., Kowalik, K., Maciejewski, G., Nowak, G., García-Hernández, D. A., Deka, B., & Adamczyk, M. 2015, Astron. & Astrophys., 573, A36

O'Hagan, A. 1978, J. R. Stat. Soc. B., 40, 1

Perryman, M., ed. 2014, The exoplanet handbook (Cambridge, UK: Cambridge University Press)

Peters, P. C. & Mathews, J. 1963, Phys. Rev, 131, 435

Pitkin, M. 2012, Mon. Not. R. Astron. Soc., 425, 2688

Preto, M., Berentzen, I., Berczik, P., & Spurzem, R. 2011, Astrophys. J, 732, L26

Pshirkov, M. S., Baskaran, D., & Postnov, K. A. 2010, Mon. Not. R. Astron. Soc., 402, 417

Queloz, D., Henry, G. W., Sivan, J. P., Baliunas, S. L., Beuzit, J. L., Donahue, R. A., Mayor, M., Naef, D., Perrier, C., & Udry, S. 2001, Astron. & Astrophys., 379, 279

Rasmussen, C. E. & Williams, C. K. I. 2006, Gaussian Processes for Machine Learning, Adaptive Computation and Machine Learning (Cambridge, MA: The MIT Press)

Ravi, V., Wyithe, J. S. B., Hobbs, G., Shannon, R. M., Manchester, R. N., Yardley, D. R. B., & Keith, M. J. 2012, Astrophys. J, 761, 84

Robert, C. P. & Casella, G. 2004, Monte Carlo Statistical Methods, 2nd edn., Springer Series in Statistics (New York, NY: Springer)

Roedig, C., Dotti, M., Sesana, A., Cuadra, J., & Colpi, M. 2011, Mon. Not. R. Astron. Soc., 415, 3033

Roques, F. et al. 2014, The extrasolar planets encyclopaedia, http://exoplanet.eu/catalog/?f=

Sato, B., Ando, H., Kambe, E., Takeda, Y., Izumiura, H., Masuda, S., Watanabe, E., Noguchi, K., Wada, S., Okada, N., Koyano, H., Maehara, H., Norimoto, Y., Okada, T., Shimizu, Y., Uraguchi, F., Yanagisawa, K., & Yoshida, M. 2003, Astrophys. J, 597, L157

Sazhin, M. V. 1978, Sov. Astron., 22, 36

Seager, S., ed. 2010, Exoplanets (Tucson, AZ: the University of Arizona Press)

Sesana, A. 2010, Astrophys. J, 719, 851

Sesana, A., Roedig, C., Reynolds, M. T., & Dotti, M. 2011, Mon. Not. R. Astron. Soc., 420, 860

Sesana, A. & Vecchio, A. 2010, Class. Quant. Grav, 27

Sesana, A., Vecchio, A., & Colacino, C. N. 2008, Mon. Not. R. Astron. Soc., 390, 192

Sesana, A., Vecchio, A., & Volonteri, M. 2009, Mon. Not. R. Astron. Soc., 394, 2255

Seto, N. 2009, Mon. Not. R. Astron. Soc., 400, L38

Shafer, G. 1982, J. Am. Stat. Assoc., 77, 325

Shannon, R. M. & Cordes, J. M. 2010, Astrophys. J, 725, 1607

Siemens, X., Mandic, V., & Creighton, J. 2007, Phys. Rev. Lett, 98, 111101

Smits, R., Kramer, M., Stappers, B., Lorimer, D. R., Cordes, J., & Faulkner, A. 2009, Astron. & Astrophys., 493, 1161

Smits, R., Tingay, S. J., Wex, N., Kramer, M., & Stappers, B. 2011, Astron. & Astrophys., 528, 108

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & Linde, A. V. D. 2002, J. R. Stat. Soc. B., Part 4, 64, 583

Stein, M. L. 1999, Interpolation of Spatial Data, Springer Series in Statistics (New York, NY: Springer)

Sudderth, E. B. 2006, PhD thesis, Masschusetts Institute of Technology, Cambridge, MA

Sudou, H., Iguchi, S., Murata, Y., & Taniguchi, Y. 2003, Science, 300, 1263

Summerscales, T. Z., Burrows, A., Finn, L. S., & Ott, C. D. 2008, Astrophys. J, 678, 1142

Taylor, J. H., Fowler, L. A., & McCulloch, P. M. 1979, Nature, 277, 437

Taylor, S. R. & Gair, J. R. 2013, Phys. Rev. D, 88, 084001

Tull, R. G. 1998, Proc. SPIE, 3355, 387

van Haasteren, R., Levin, Y., McDonald, P., & Lu, T. 2009, Mon. Not. R. Astron. Soc., 395, 1005

van Haasteren, R., Mingarelli, C. M. F., Vecchio, A., & Lassus, A. 2013, Analysis of the first IPTA Mock Data Challenge by the EPTA timing data analysis working group, arXiv: 1301.6673

van Haasteren, R. et al. 2011, Mon. Not. R. Astron. Soc., 414, 3117

Verbiest, J. P. W., Bailes, M., van Straten, W., Hobbs, G. B., Edwards, R. T., Manchester, R. N., Bhat, N. D. R., Sarkissian, J. M., Jacoby, B. A., & Kulkarni, S. R. 2008, Astrophys. J, 679, 675

Walker, S. G. 2010, in Bayesian Nonparametrics, ed. N. L. Hjort, C. Holmes, P. Müller, & S. G. Walker, Cambridge Series in Statistical and Probabilistic Mathematics (Cambridge, UK: Cambridge University Press)

Winn, J. N. 2010, in Exoplanets, ed. S. Seager (Tucson, AZ: the University of Arizona Press)

Yardley, D. R. B., Hobbs, G. B., Jenet, F. A., Verbiest, J. P. W., Wen, Z. L., Manchester, R. N., Coles, W. A., van Straten, W., Bailes, M., Bhat, N. D. R., Burke-Spolaor, S., Champion, D. J., Hotan, A. W., & Sarkissian, J. M. 2010, Mon. Not. R. Astron. Soc., 407, 669

Zieliński, P., Niedzielski, A., Wolszczan, A., Adamów, M., & Nowak, G. 2012, Astron. & Astrophys., 547, A91

Vita

XIHAO DENG

(814) · 321 · 3693 ◇ xud104@psu.edu
Davey 104 ◇ University Park, PA 16802

EDUCATION

The Pennsylvania State University *August 2008 - May 2015*
Ph.D candidate in physics
Doctor Advisor: Lee Samuel Finn
Overall GPA: 3.91/4.0

Beijing Normal University, China *August 2004 - June 2008*
Bachelor of Science in Physics

RESEARCH EXPERIENCE

Developing gravitational wave detection methodology with pulsar timing arrays.

Developing detection methodology of exoplanetary systems around giant stars with radial velocity observations.

AWARDS

2008-2009 University Graduate Fellowship, Pennsylvania State University

2009-2012 Duncan Fellowship, Pennsylvania State University

2011-2012 Downsbrough Graduate Fellowship, Pennsylvania State University

2012-2013 Mebus Gaduate Fellowship, Pennsylvania State University

2012-2013 Eklund Lecture Award, Pennsylvania State University

2013-2014 W. Donald Miller Graduate Fellowship, Pennsylvania State University

SELECTED PUBLICATION

1. X. Deng, L. S. Finn, Monthly Notices of the Royal Astronomical Society, 414, 50-58, 2011
2. X. Deng, American Institute of Physics Conference Proceedings, 1467, 231-234, 2012
3. X. Deng, Physical Review D, 90, 024020, 2014
4. X. Deng, Physical Review D, 90, 104029, 2014.