Mon. Not. R. Astron. Soc. **395,** 1005–1014 (2009) doi:10.1111/j.1365-2966.2009.14590.x

On measuring the gravitational-wave background using Pulsar Timing Arrays

Rutger van Haasteren,[1⋆] Yuri Levin,[1,2] Patrick McDonald[3] and Tingting Lu[3,4]

[1]*Leiden University, Leiden Observatory, PO Box 9513, NL-2300 RA Leiden, the Netherlands*
[2]*Leiden University, Lorentz Institute, PO Box 9506, NL-2300 RA Leiden, the Netherlands*
[3]*CITA, 60 St. George Street, Toronto, Ontario M5S 3H8, Canada*
[4]*Department of Astronomy, University of Toronto, 60 St. George Street, Toronto, Ontario M5S 3H8, Canada*

Accepted 2009 February 3. Received 2009 February 3; in original form 2008 November 28

ABSTRACT

The long-term precise timing of Galactic millisecond pulsars holds great promise for measuring the long-period (months to years) astrophysical gravitational waves. Several gravitational-wave observational programs, called Pulsar Timing Arrays (PTA), are being pursued around the world.

Here, we develop a Bayesian algorithm for measuring the stochastic gravitational-wave background (GWB) from the PTA data. Our algorithm has several strengths: (i) it analyses the data without any loss of information; (ii) it trivially removes systematic errors of known functional form, including quadratic pulsar spin-down, annual modulations and jumps due to a change of equipment; (iii) it measures simultaneously both the amplitude and the slope of the GWB spectrum and (iv) it can deal with unevenly sampled data and coloured pulsar noise spectra. We sample the likelihood function using Markov Chain Monte Carlo simulations. We extensively test our approach on mock PTA data sets and find that the algorithm has significant benefits over currently proposed counterparts. We show the importance of characterizing all red noise components in pulsar timing noise by demonstrating that the presence of a red component would significantly hinder the detection of the GWB.

Lastly, we explore the dependence of the signal-to-noise ratio on the duration of the experiment, number of monitored pulsars and the magnitude of the pulsar timing noise. These parameter studies will help formulate observing strategies for the PTA experiments.

Key words: gravitational waves – methods: data analysis – pulsars: general.

1 INTRODUCTION

At the time of this writing several large projects are being pursued in order to directly detect astrophysical gravitational waves (GWs). This paper concerns a program to detect GWs using pulsars as nearly perfect Einstein clocks. The practical idea is to time a set of millisecond pulsars (called the 'Pulsar Timing Array' or PTA) over a number of years (Foster & Backer 1990). Some of the millisecond pulsars create pulse trains of exceptional regularity. By perturbing the space–time between a pulsar and the Earth, the GWs will cause extra deviations from the periodicity in the pulse arrival times (Estabrook & Wahlquist 1975; Sazhin 1978; Detweiler 1979). Thus from the measurements of these deviations (called 'timing residuals' or TR), one may measure the GWs. Currently, several PTA project are operating around the globe. First, at the Arecibo Radio Telescope in North America several millisecond pulsars have been timed for a number of years. These observations have already been used to place interesting upper limits on the intensity of GWs which are passing through the Galaxy (Kaspi, Taylor & Ryba 1994; Lommen 2001). Together with the Green Bank Telescope, the Arecibo Radio Telescope will be used as an instrument of North American Nanohertz Observatory for Gravitational Waves (NANOGrav), the North American PTA. Secondly, the European PTA is being set up as an international collaboration between Great Britain, France, Netherlands, Germany and Italy and will use five European radio telescopes to monitor about 20 millisecond pulsars (Stappers et al. 2006). Finally, the Parkes PTA in Australia has been using the Parkes multibeam radio telescope to monitor 20 millisecond pulsars (Manchester 2006). Some of the Parkes and Arecibo data have also been used to place the most stringent limits on the gravitational-wave background (GWB) to date (Jenet et al. 2006).

One of the main astrophysical targets of the PTAs is the stochastic background of the GWs (GWB). This GWB is thought to be generated by a large number of black-hole binaries which are thought to

⋆E-mail: haasteren@strw.leidenuniv.nl

be located at the centres of galaxies (Begelman, Blandford & Rees 1980; Phinney 2001; Jaffe & Backer 2003; Wyithe & Loeb 2003; Sesana et al. 2005), by relic GWs (Grishchuk 2005) or, more speculatively, by cusps in the cosmic-string loops (Damour & Vilenkin 2005). This paper develops an algorithm for the optimal PTA measurement of such a GWB.

The main difficulty of such a measurement is that not only GWs create the pulsar TRs. Irregularities of the pulsar-beam rotation (called the 'timing noise'), the receiver noise, the imprecision of local clocks, the polarization calibration of the telescope (Britton 2000) and the variation in the refractive index of the interstellar medium all contribute significantly to the TRs making it a challenge to separate these noise sources from the GW signal. However, the GWB is expected to induce correlations between the TRs of different pulsars. These correlations are of a specific functional form (given by equation 9), which is different from those introduced by other noise sources (Hellings & Downs 1983). Jenet, Hobbs, Lee & Manchester (2005, hereafter J05) have invented a clever algorithm which uses the uniqueness of the GWB-induced correlations to separate the GWB from other noise sources and thus to measure the magnitude of the GWB. Their idea was to measure the TR correlations for all pairs of the PTA pulsars and check how these correlations depend on the sky angles between the pulsar pairs. J05 have derived a statistic which is sensitive to the functional form of the GWB-induced correlation; by measuring the value of this statistic, one can infer the strength of the GWB. While J05 algorithm appears robust, we believe that in its current form it does have some drawbacks, in particular.

(i) The statistic used by J05 is non-linear and non-quadratic in the pulsar TRs, which makes its statistical properties non-transparent.

(ii) The pulsar pairs with the high and low intrinsic timing noise make equal contributions to the J05 statistic, which is clearly not optimal.

(iii) The J05 statistic assumes that the TRs of all the PTA pulsars are measured during each observing run, which is generally not the case.

(iv) The J05 signal-to-noise (S/N) ratio analysis relies on the prior knowledge of the intrinsic timing noise, and there is no clean way to separate this timing noise from the GWB.

(v) The prior spectral information on GWB is used for whitening the signal; however, there is no proof that this is an optimal procedure. The spectral slope of the GWB is not measured.

In this paper, we develop an algorithm which addresses all of the problems outlined above. Our method is based on essentially the same idea as that of J05: we use the unique character of the GWB-induced correlations to measure the intensity of the GWB. The algorithm we develop below is Bayesian, and by construction it uses optimally all of the available information. Moreover, it deals correctly and efficiently with all systematic contributions to the TRs which have a known functional form, i.e. the quadratic pulsar spin-downs, annual variations, one-time discontinuities (jumps) due to equipment change, etc. Many parameters of the timing model (the model popular pulsar timing packages use to generate TRs from pulsar arrival times) fall in this category.

The plan of this paper is as follows. In the next section, we review the theory of the GWB-generated TRs and introduce our model for other contributions to the TRs. In Section 3, we explain the principle of the Bayesian analysis for GWB measurement with a PTA, and we evaluate the Bayesian likelihood function. There, we also show how to analytically marginalize over the contributions of known functional form but unknown amplitude (i.e. annual variations, quadratic residuals due to pulsar spin-down, etc.). The details of this calculation are laid out in Appendix A. Section 4 discusses the numerical integration technique which we use in our likelihood analysis: the Markov Chain Monte Carlo (MCMC). In Section 5, we show the analyses of mock PTA data sets. For each mock data set, we construct the probability distribution for the intensity of the GWB and demonstrate its consistency with the input mock data parameters. We study the sensitivity of our algorithm for different PTA configurations and investigate the dependence of the S/N ratio on the duration of the experiment, on redness and magnitude of the pulsar timing noise and on the number of clocked pulsars. In Section 6, we summarize our results.

2 THE THEORY OF GW-GENERATED TIMING RESIDUALS

2.1 Timing residual correlation

The measured millisecond pulsar TRs contain contributions from several stochastic and deterministic processes. The latter include the gradual deceleration of the pulsar spin, resulting in a pulsar rotational period derivative which induces TRs varying quadratically with time (hereafter referred to as 'quadratic spin-down'), the annual variations due to the imperfect knowledge of the pulsar positions on the sky, the ephemeris variations caused by the known planets in the Solar system and the jumps due to equipment change (Manchester 2006). The stochastic component of the TRs will be caused by the receiver noise, clock noise, intrinsic timing noise, the refractive index fluctuations in the interstellar medium and, most importantly for us, the GWB. For the purposes of this paper, we restrict ourselves to considering the quadratic spin-downs, intrinsic timing noise and the GWB; other components can be similarly included, but we omit them for mathematical simplicity. In this case, the ith TR of the ath pulsar can be written as

$$\delta t_{ai} = \delta t_{ai}^{\mathrm{GW}} + \delta t_{ai}^{\mathrm{PN}} + Q(t_{ai}), \tag{1}$$

where $\delta t_{ai}^{\mathrm{GW}}$ and $\delta t_{ai}^{\mathrm{PN}}$ are caused by the GWB and the pulsar timing noise, respectively, and

$$Q_a(t_{ai}) = A_{a1} + A_{a2}t_{ai} + A_{a3}t_{ai}^2 \tag{2}$$

represent the quadratic spin-down. One expects the timing noise from different pulsars to be uncorrelated, while the GWB will cause correlations in the TRs between different pulsars. Therefore, the information about GWB can be extracted by correlating the TR data between the different pulsars (J05). If one assumes that both GWB-generated residuals and the intrinsic timing noise are stochastic Gaussian processes, then we can represent them by the $(n \times n)$ coherence matrices:

$$\begin{aligned} \left\langle \delta t_{ai}^{\mathrm{GW}} \delta t_{bj}^{\mathrm{GW}} \right\rangle &= \mathbf{C}_{(ai)(bj)}^{\mathrm{GW}}, \\ \left\langle \delta t_{ai}^{\mathrm{PN}} \delta t_{bj}^{\mathrm{PN}} \right\rangle &= \mathbf{C}_{(ai)(bj)}^{\mathrm{PN}}, \end{aligned} \tag{3}$$

with the total coherence matrix given by

$$\mathbf{C}_{(ai)(bj)} = \mathbf{C}_{(ai)(bj)}^{\mathrm{GW}} + \mathbf{C}_{(ai)(bj)}^{\mathrm{PN}}. \tag{4}$$

The TRs are then distributed as a multidimensional Gaussian:

$$P(\delta \boldsymbol{t}) = \frac{1}{\sqrt{(2\pi)^n \det \mathbf{C}}} \exp\left\{ -\frac{1}{2} \sum_{(ai)(bj)} [\delta \boldsymbol{t}_{(ai)} - Q_a(t_{ai})] \mathbf{C}_{(ai)(bj)}^{-1} [\delta \boldsymbol{t}_{(bj)} - Q_b(t_{bj})] \right\}, \tag{5}$$

where P denotes the probability distribution of the TRs. To be able to use equation (5) we must

(i) be able to evaluate the GWB-induced coherence matrix from the theory, as a function of variables that parametrize the GWB spectrum and

(ii) introduce well-motivated parametrization of the pulsar timing noise. In this work, the spectral density of the stochastic GW background is taken to be a power law (Maggiore 2000; Phinney 2001; Jaffe & Backer 2003; Wyithe & Loeb 2003)

$$S_h = A^2 \left(\frac{f}{\mathrm{yr}^{-1}}\right)^{-\gamma}, \tag{6}$$

where S_h represents the spectral density, A is the GW amplitude, f is the GW frequency and γ is an exponent characterizing the GWB spectrum. If the GWB is dominated by the supermassive black hole binaries, then $\gamma = 7/3$ (Phinney 2001). This definition is equivalent to the use of the characteristic strain as defined in Jenet et al. (2006):

$$h_c = A \left(\frac{f}{\mathrm{yr}^{-1}}\right)^{\alpha}, \tag{7}$$

with $\gamma = 1 - 2\alpha$. The GWB-induced coherence matrix is then given by

$$\mathbf{C}^{\mathrm{GW}}_{(ai)(bj)} = \frac{A^2 \alpha_{ab}}{(2\pi)^2 f_{\mathrm{L}}^{1+\gamma}} \left\{ \Gamma(-1-\gamma) \sin\left(\frac{-\pi\gamma}{2}\right) (f_{\mathrm{L}}\tau)^{\gamma+1} - \sum_{n=0}^{\infty} (-1)^n \frac{(f_{\mathrm{L}}\tau)^{2n}}{(2n)!(2n-1-\gamma)} \right\}. \tag{8}$$

Here, α_{ab} is the geometric factor given by

$$\alpha_{ab} = \frac{3}{2} \frac{1-\cos\theta_{ab}}{2} \ln\left(\frac{1-\cos\theta_{ab}}{2}\right) - \frac{1}{4} \frac{1-\cos\theta_{ab}}{2} + \frac{1}{2} + \frac{1}{2}\delta_{ab}, \tag{9}$$

where θ_{ab} is the angle between pulsars a and b (Hellings & Downs 1983), $\tau = 2\pi(t_{ai} - t_{bj})$, Γ is the gamma function and f_{L} is the low cut-off frequency chosen so that $1/f_{\mathrm{L}}$ is much greater than the duration of the PTA operation. Introducing f_{L} is a mathematical necessity, since otherwise the GWB-induced correlation function would diverge. However, we show below that the low-frequency part of the GWB is indistinguishable from an extra spin-down of all pulsars which we already correct for and that our results do not depend on the choice of f_{L} provided that $f_{\mathrm{L}}\,\tau \ll 1$.

The pulsar timing noise is assumed to be Gaussian, with a certain functional form of the power spectrum. The true profile of the millisecond pulsar timing noise spectrum is not well known at present time. The TRs of the most precisely observed pulsars indicate that pulsar timing noise has a white and poorly constrained red component (Verbiest & Hobbs, private communications).

For the purposes of this paper, we will always choose the spectra to be of the same functional form for all pulsars, but this is not an inherent limitation of the algorithm. We consider three cases of pulsar timing noise spectra:

(i) white (flat) spectra;
(ii) Lorentzian spectra and
(iii) power-law spectra.

Obviously, one could also consider a timing noise which is a superposition of these components; we do not do this at this exploratory stage. If we choose the pulsar timing noise spectrum to be white, with an amplitude N_a, the resulting correlation matrix becomes:

$$\mathbf{C}^{\mathrm{PN-white}}_{(ai)(bj)} = N_a^2 \delta_{ab} \delta_{ij}. \tag{10}$$

The Lorentzian spectrum is a red spectrum with a typical frequency that determines the redness of the timing noise:

$$S_a(f) = \frac{N_a^2}{f_0 \left[1 + \left(\frac{f}{f_0}\right)^2\right]}, \tag{11}$$

which yields the following correlation matrix:

$$\mathbf{C}^{\mathrm{PN-lor}}_{(ai)(bj)} = N_a^2 \delta_{ab} \exp(-f_0 \tau), \tag{12}$$

where f_0 is a typical frequency and N_a is the amplitude.

By using a power-law spectral density with amplitude N_a and spectral index γ_a, one gets a timing noise coherence matrix analogous to the one in equation (8):

$$\mathbf{C}^{\mathrm{PN-pl}}_{(ai)(bj)} = \frac{N_a^2 \delta_{ab}}{f_{\mathrm{L}}^{\gamma_a - 1}} \left\{ \Gamma(1-\gamma_a) \sin\left(\frac{\pi\gamma_a}{2}\right) (f_{\mathrm{L}}\tau)^{\gamma_a - 1} - \sum_{n=0}^{\infty} (-1)^n \frac{(f_{\mathrm{L}}\tau)^{2n}}{(2n)!\,(2n+1-\gamma_a)} \right\}. \tag{13}$$

3 BAYESIAN APPROACH

3.1 Basic ideas

The method described in this report is based upon a Bayesian approach to the parameter inference. The general idea of the method is to (i) assume that the physical processes which produce the TRs can be characterized by several parameters and (ii) use the Bayes theorem to derive from the measured data the probability distribution of the parameters of our interest. In our case, we assume that the TRs are created by

(i) the GWB; we parametrize it by its amplitude A and slope γ, as in equation (6),

(ii) the intrinsic timing noise of the 20 monitored millisecond pulsars. We assume that the timing noise of each of the pulsars is the random Gaussian noise, with a variety of possible spectra described in the previous section. We will refer to the variables parametrizing the timing noise spectral shape as TN_a and

(iii) the quadratic spin-downs, parametrized for each of the pulsars by A_{a1}, A_{a2} and A_{a3}, cf. equation (2).

With these assumptions, we will write down below the expression for the probability distribution P(data|parameters) of the data, as a function of the parameters. By the Bayes theorem, we can then compute the posterior distribution function; the probability distribution of the parameters given a certain data set:

$$P(\text{parameters}|\text{data}) = P(\text{data}|\text{parameters}) \times \frac{P_0(\text{parameters})}{P(\text{data})}. \tag{14}$$

Here, P_0(parameters) is the prior probability of the unknown parameters, which represents all our current knowledge about these parameters and P(data) is the Bayesian evidence, which we will use here as a normalization factor to ensure that $P(A, \gamma, TN_a, A_{a1}, A_{a2}, A_{a3}|\text{data})$ integrates to unity over the parameter space. We note here that the Bayesian evidence is in essence a goodness-of-fit measure that can be used for model selection. However, we will ignore the Bayesian evidence in this work and postpone the model selection part of the algorithm to future work. For our purposes, we are only interested in A and γ, which means that we have to integrate $P(A, \gamma, TN_a, A_{a1}, A_{a2}, A_{a3}|\text{data})$ over all of the other parameters. Luckily, as we show below, for a uniform prior the integration over A_{a1},

A_{a2} and A_{a3} can be performed analytically. This amounts to the *removal* of the quadratic spin-down component to the pulsar data. We emphasize that this removal technique is quite general and can be readily applied to unwanted signal of any known functional form (i.e. annual modulations, jumps, etc. – see Section 3.2), even if those parameters have already been fit for while calculating the TRs. The integration over TN_a must be performed numerically.

In this work, we will use MCMC simulation as a multidimensional integration technique. Besides flat priors for most of the parameters, we will use slightly peaked priors for parameters which have non-normalizable likelihood functions. This ensures that the Markov Chain can converge.

In the rest of this paper, we detail the implementation and tests of our algorithm.

3.2 Removal of the quadratic spin-downs and other systematic signals of known functional form

While this section is written with the PTA in mind, it may well be useful for other applications in pulsar astronomy. We thus begin with a fairly general discussion and then make it more specific for the PTA case.

Consider a random Gaussian process δx_i^{G} with a coherence matrix $\mathbf{C}(\sigma)$, which is contaminated by several systematic signals with known functional forms $f_p(t_i)$ but a priori unknown amplitudes ξ_p. Here, σ is a set of interesting parameters which we want to determine from the data δx. The resulting signal is given by

$$\delta x_i = \delta x_i^{\mathrm{G}} + \sum_p \xi_p f_p(t_i), \tag{15}$$

or, in the vector form, by

$$\delta \boldsymbol{x} = \delta \boldsymbol{x}^{\mathrm{G}} + \mathbf{F}\boldsymbol{\xi}. \tag{16}$$

Here, the components of the vectors $\delta\boldsymbol{x}$, $\delta\boldsymbol{x}^{\mathrm{G}}$ and $\boldsymbol{\xi}$ are given by δx_i, δx_i^{G} and ξ_p, respectively, and $\mathbf{F}$ is the non-square matrix with the elements $F_{ip} = f_p(t_i)$. Note that the dimensions of $\delta\boldsymbol{x}$ and $\boldsymbol{\xi}$ are different. The Bayesian probability distribution for the parameters is given by

$$P(\sigma, \boldsymbol{\xi}|\delta\boldsymbol{x}) = \frac{M}{\sqrt{\det \mathbf{C}}} \exp\left[-\frac{1}{2}(\delta\boldsymbol{x} - \mathbf{F}\boldsymbol{\xi})\mathbf{C}^{-1}(\delta\boldsymbol{x} - \mathbf{F}\boldsymbol{\xi})\right] \times P_0(\sigma, \boldsymbol{\xi}), \tag{17}$$

where P_0 is the prior probability and M is the normalization. Since we are only interested in σ, we can integrate $P(\sigma, \boldsymbol{\xi}|\delta\boldsymbol{x})$ over the variables $\boldsymbol{\xi}$. This process is referred to as marginalization; it can be done analytically if we assume a flat prior for $\boldsymbol{\xi}$ [i.e. if $P_0(\sigma, \boldsymbol{\xi})$ is $\boldsymbol{\xi}$-independent], since ξ_p enter at most quadratically into the exponential above. After some straightforward mathematics which we have detailed in Appendix A, we get

$$P(\sigma|\delta x) = \frac{M'}{\sqrt{\det(\mathbf{C})\det(\mathbf{F}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{F})}} \times \exp\left[-\frac{1}{2}\delta\boldsymbol{x}\mathbf{C}'\delta\boldsymbol{x}\right], \tag{18}$$

where M' is the normalization and

$$\mathbf{C}' = \mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{F}(\mathbf{F}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{F})^{-1}\mathbf{F}^{\mathrm{T}}\mathbf{C}^{-1}, \tag{19}$$

and the T superscript stands for the transposed matrix. Equation (18) is one of the main equations of this paper, since it provides a statistically rigorous way to remove (i.e. marginalize over) the unwanted systematic signals from random Gaussian processes. One can check directly that the above expression for $P(\sigma|\delta\boldsymbol{x})$ is insensitive to the values ξ_p of the amplitudes of the systematic signals in the equation (15).

We now apply this formalism to account for the quadratic spin-downs in the PTA. As in Section 2, it will be convenient to use the two-index notation for the TRs, δt_{ai} measured at the time t_{ai}, where a is the pulsar index and i is the number of the TR measurement for pulsar a. The space of the spin-down parameters A_{aj}, $j = 1, 2, 3$ has $3N$ dimensions, where N is the number of pulsars in the array. In the component language, we write

$$\delta t_{ai} = \delta t_{ai}^{\mathrm{G}} + \sum_{b,j} \mathbf{F}_{(ai)(bj)} A_{bj}, \tag{20}$$

where

$$\mathbf{F}_{(ai)(bj)} = \delta_{ab} t_{ai}^{j-1}, \tag{21}$$

δt^{G} is the part of the TR due to a random Gaussian process (i.e. GWB, timing noise, etc.) and $j = 1, 2, 3$. The quantities $F_{(ai)(bj)}$ are components of the matrix operator which acts on the $3N$-dimensional vector in the parameter space and produces a vector in the TR space. For example, for 20 pulsars, each with 250 TR observations, the matrix $\mathbf{F}_{(ai)(bj)}$ has $20 \times 250 = 5000$ rows, each marked by two indices $a = 1, \ldots, 20$, $i = 1, \ldots, 250$, and $20 \times 3 = 60$ columns, each marked by two indices $b = 1, \ldots, 20$, $j = 1, 2, 3$. Thus in the vector form, one can write equation (20) as

$$\delta \boldsymbol{t} = \delta \boldsymbol{t}^{\mathrm{G}} + \mathbf{F}\boldsymbol{A}, \tag{22}$$

which is identical to the equation (16). We thus can use equation (18) to remove the quadratic spin-down contribution from the PTA data.

Although we only demonstrate this technique for quadratic spin-down, this removal technique will be useful for treating other noise sources in the PTA. All sources of which the functional form is known (and therefore can be fit for, as most popular pulsar timing packages do) can be dealt with, i.e.

(i) The annual variation of the TRs due to the imprecise knowledge of the pulsar position on the sky. The annual variation in each of the pulsars will be a predictable function of the associated two small angular errors (latitude and longitude). Thus, our parameter space will expand by $2N$, but this will still keep the $\mathbf{F}$ matrix manageable.

(ii) Changes of equipment will introduce extra jumps, and must be taken into account. This is trivial to deal with using the techniques described above.

(iii) Some of the millisecond pulsars are in binaries, and their orbital motion must be subtracted. The errors one makes in these subtractions will affect the TRs. They can be parametrized and dealt with using the techniques of this section (we thank Jason Hessels for pointing this out).

3.3 Low-frequency cut-off

All predictions for GWB spectrum show a steep power law $\propto f^{-\gamma}$, where for black hole binaries $\gamma = 7/3$ (Phinney 2001). Physically, there is a low-frequency cut-off to the spectrum, due to the fact that black-hole binaries with periods greater than 1000 years shrink mostly due to the external friction (i.e. scattering of circumbinary stars or excitation of density waves in a circumbinary gas disc), and not to gravitational radiation. However, while the exact value of the low-frequency cut-off is poorly constrained, the PTA should not be sensitive to it since the duration of the currently planned experiments is much shorter than 1000 years. In this section, we show this formally by explicitly introducing the low-frequency cut-off and by demonstrating that our Bayesian probabilities are insensitive to its value.

Consider the expression in equation (8) for the GWB-generated correlation matrix for the TRs. This expression contains an integral of the form

$$I = \int_{f_{\rm L}}^{\infty} \cos(f\tau) f^{-(\gamma+2)}\,{\rm d}f, \quad (23)$$

where $\tau = 2\pi(t_i - t_j)$. When the low-frequency cut-off is much smaller than the inverse of the experiment duration, i.e. when $f_{\rm L}\tau \ll 1$, the integral above can be expanded as

$$I = B\tau^{\gamma+1} + \frac{1}{f_{\rm L}^{\gamma+1}}\left\{\frac{1}{(\gamma+1)} - \frac{(f_{\rm L}\tau)^2}{2(\gamma-1)} + {\rm O}[(f_{\rm L}\tau)^4]\right\}, \quad (24)$$

where

$$B = \Gamma(-1-\gamma)\sin\left(\frac{-\pi\gamma}{2}\right)\tau^{\gamma+1}. \quad (25)$$

In the expansion above, we have assumed $1 < \gamma < 3$. The terms which contain $f_{\rm L}$ diverge when $f_{\rm L}$ goes to zero and scale as τ^0 or τ^2 with respect to the time interval. We now show that these divergent terms get absorbed in the process of elimination of the quadratic spin-downs.

Suppose that we add to the TRs of a pulsar a quadratic spin-down term, $A_1 + A_2 t + A_3 t^2$. The spin-down-removal procedure described in the previous section makes our results completely insensitive to this addition: As could be arbitrarily large but the measured GWB would still be the same. Clearly, the same is true if one treats A_1, A_2, A_3 not as fixed numbers, but as random variables drawn from some Gaussian distribution. The correlation introduced into the TRs by adding a *random* quadratic spin-down is given by

$$\begin{aligned}\langle\delta t_i \delta t_j\rangle = &\left\langle A_1^2\right\rangle + \langle A_1 A_2\rangle(t_i + t_j)\\ &+ 2\left\langle A_2^2\right\rangle t_i t_j + \langle A_1 A_3\rangle\left(t_i^2 + t_j^2\right)\\ &+ \langle A_2 A_3\rangle t_i t_j (t_i + t_j) + \left\langle A_3^2\right\rangle t_i^2 t_j^2.\end{aligned} \quad (26)$$

The $f_{\rm L}$-dependent part of equation (24) contains terms which scale as $t_i^2 + t_j^2$, $t_i t_j$ and const, and thus have the same functional t_i, t_j dependence as some of the terms in equation (26). Since the terms in equation (26) can be made arbitrarily large, it is clear that the terms corresponding to the low-frequency cut-off could be absorbed into the correlation function corresponding to the quadratic spin-down with the stochastic coefficients. We have made this argument for the TRs from a single pulsar, but it is trivial to extend it to the case of multiple pulsars. Thus, our results are not sensitive to the actual choice of the $f_{\rm L}$ so long as $f_{\rm L}\tau \ll 1$; this is confirmed by direct numerical tests.

4 NUMERICAL INTEGRATION TECHNIQUES

4.1 Metropolis Monte Carlo

The Bayesian probability distribution for the PTA is computed in multidimensional parameter space, where all of the parameters except 2 characterize intrinsic pulsar timing noise and other potential interferences. To obtain meaningful information about the GWB, we need to integrate the probability function over all of the unwanted parameters. This is a challenging numerical task: a direct numerical integration over more than several parameters is prohibitively computationally expensive. Fortunately, numerical shortcuts do exist and the most common among them is the MCMC simulation. In a typical MCMC, a set of semirandom walkers samples the parameter space in a clever way, each generating a large number of sequential locations called a *chain* (Newman & Barkema 1999). After a sufficient number of steps, the density of points of the chain becomes proportional to the Bayesian probability distribution. The number of steps required for the chain convergence scales linearly with the number of dimensions of the parameter space; typically, few times 10^4 steps are required for reliable convergence. In this paper, we use the Metropolis (Newman & Barkema 1999) algorithm for generating the chain, which can be used with an arbitrary distribution, the proposal distribution, for generating new locations of the chain. We use a Gaussian proposal distribution, centred at the current location in the parameter space. During an initial period, the burn-in period, the width of the proposal distribution in all dimensional directions is set to yield the asymptotically optimal acceptance rate of 23.4 per cent for the Metropolis algorithm (Roberts, Gelman & Gilks 1997). At the end of the MCMC simulation, we check the convergence of the chain using the bootstrap method (Efron 1979). We also calculate the global maximum likelihood value for all parameters using a conjugate directions search (Brent 1973).

4.2 Current MCMC computational cost

The greatest computational challenge in constructing the chain is the fast evaluation of the matrix $\mathbf{C}^{-1}$ in equations (18) and (19). If 250 TRs are measured for each of the pulsars (50 weeks for 5 years), the size of the matrix $\mathbf{C}$ becomes (5000×5000). We find it takes about 20 s to invert $\mathbf{C}$ and thus about 1.5 times as much to arrive to the next point in the chain. Therefore, for the required 10^5 chain points to get the convergent distribution, we need of the order of 1 month of the single-processor computational time. On a cluster, this can be done in a couple of days. We emphasize that this is an order of n^3 process. For matrices of (2000×2000), the calculation can be done overnight on a single modern workstation, but for ($10^4 \times 10^4$) the calculation is already a serious challenge.

For the currently projected size of the data sets (Manchester 2006), the amount of TRs will most likely not exceed the 250 (Hobbs, private communications). Thus, the brute-force method presented here is not computationally expensive for the projected data volume over the next 5 years.

4.3 Choosing a suitable prior distribution

For some models (e.g. the power-law spectal density for pulsar timing noise), the likelihood function proves to be not normalizable. This would pose a serious problem in combination with uniform priors as the nuisance parameters then cannot be marginalized and the posterior cannot represent a probability distribution. Although this is a sign that our model is incorrect (infinite Bayesian evidence/normalization), this can be easily solved with a different parametrization. We can always change coordinates in parameter space to a set for which all parameters have a finite domain, which guarantees that our likelihood function is normalizable. This procedure is equivalent to choosing a different prior (the Jacobian in the case of a coordinate transformation) for the original set. We therefore argue that we need to choose an appropriate prior for the non-normalizable parameters. We propose to use a Lorentzian-shaped profile:

$$P_0(\gamma_i) = \frac{\Delta_i}{\pi\left(\Delta_i^2 + \gamma_i^2\right)}, \quad (27)$$

where γ_i is the parameter for which we are constructing a prior, and Δ_i is some typical width/value for that parameter.

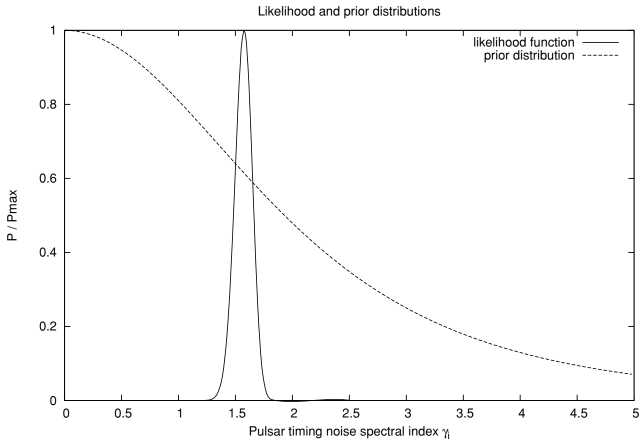

Figure 1. The likelihood and prior distribution for a pulsar timing noise spectral index parameters γ_i. The solid line represents the likelihood function. It is sharply peaked and it looks as if it drops to zero for high γ_i. However, for high γ_i it will have a constant non-negligible value. The dashed line represents our chosen prior distribution. The prior is normalizable, and its application makes the posterior distribution normalizable as well.

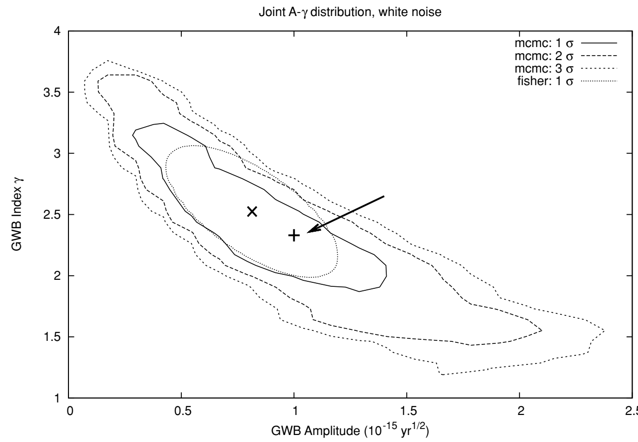

Figure 2. The GW likelihood function (GW amplitude, GW slope versus probability density contours), determined with the MCMC method for a set of mock data with 20 pulsars, and 100 data points per pulsar approximately evenly distributed over 5 years. Each pulsar has a white timing noise of 100 ns. The true GW amplitude and slope are shown as a '+'with an arrow and the maximum likelihood values are shown as 'x'. The contours are in steps of σ, with the inner one at 1σ. The 1σ contour of the Gaussian approximation is also shown.

As an example, we show the likelihood function and the prior for the pulsar timing noise spectral index parameter of equation (13) in Fig. 1. The likelihood function seems to drop to zero for high γ_i, but it actually has a non-negligible value for all γ_i greater than the maximum likelihood value. The broadness of the prior is chosen such that it does not change the representation of the significant part of the likelihood in the posterior, but it does make sure that the posterior is normalizable.

4.4 Generating mock data

In order to generate mock data, we produce a realization of the multidimensional Gaussian process of equation (5), as follows. We rewrite equation (5) using a basis in which $\mathbf{C}$ is diagonal:

$$P(\delta \boldsymbol{t}) = \prod_{i=1}^{n} \frac{1}{\sqrt{\lambda_i}} \varphi\left(\frac{y_i}{\sqrt{\lambda_i}}\right), \tag{28}$$

where,

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right). \tag{29}$$

Here, λ_i are the eigenvalues of $\mathbf{C}$ and

$$\boldsymbol{y} = \mathbf{T}^{-1}\delta \boldsymbol{t}, \tag{30}$$

where $\mathbf{T}$ is the transformation matrix which diagonalizes $\mathbf{C}$:

$$(\mathbf{T}^{-1}\mathbf{C}\mathbf{T})_{ij} = \lambda_i \delta_{ij}. \tag{31}$$

Thus, we follow the following steps.

(i) Diagonalize matrix $\mathbf{C}$, find $\mathbf{T}$ and λ_i.
(ii) Choose y_i from random Gaussian distributions of widths $\sqrt{\lambda_i}$.
(iii) Compute the TRs via equation (30).

It is then trivial to add deterministic processes, like quadratic spin-downs, to the simulated TRs.

5 TESTS AND PARAMETER STUDIES

We test our algorithm by generating mock TRs for a number of millisecond pulsars which are positioned randomly in the sky. We found it convenient to parametrize the GWB spectrum by (cf. equation 6)

$$S_h(f) = A^2 \left(\frac{f}{\mathrm{yr}^{-1}}\right)^{-\gamma}. \tag{32}$$

Our mock TRs are a single realization of GWB for some values of A and γ and the pulsar timing noise. Random quadratic spin-down terms are added. We then perform several separate investigations as follows.

5.1 Single data set tests

Our algorithm is tested on several data sets in the following way. The mock data sets were generated with parameters resembling an experiment of 20 pulsars, with observations approximately every 5 weeks for 5 years. The pulsar timing noise was set to an optimistic level of 100 ns each (rms TRs). In all cases, the level of GWB has been set to $A = 10^{-15}\,\mathrm{yr}^{1/2}$, with $\gamma = 7/3$. This level of GWB is an order of magnitude smaller than the most recent upper limits of this type (Jenet et al. 2006). We then analyse these mock data using the MCMC method. In Figs 2–4, we see examples of the joint $A - \gamma$ probability distribution, obtained by these analyses. For each data set, we also calculate the maximum likelihood value of all parameters using a conjugate directions search. The algorithm gives results consistent with the input parameters (i.e. they recover the amplitude and the slope of the GWB within measurement errors). This was observed in all our tests.

For all data sets, we also calculated the Fisher information matrix, a matrix consisting of second-order derivatives to all parameters, at the maximum likelihood points. We can use this matrix to approximate the posterior by a multidimensional Gaussian, since for some particular models this approximation is quite good. The Fisher information matrix can be calculated in a fraction of the time needed to perform a full MCMC analysis. For all data sets, we have plotted the 1σ contour of the multidimensional Gaussian approximation.

As an extra test, we have also used data sets generated by the popular pulsar timing package TEMPO2 (Hobbs, Edwards & Manchester 2006) with a suitable GWB simulation plug-in (Hobbs

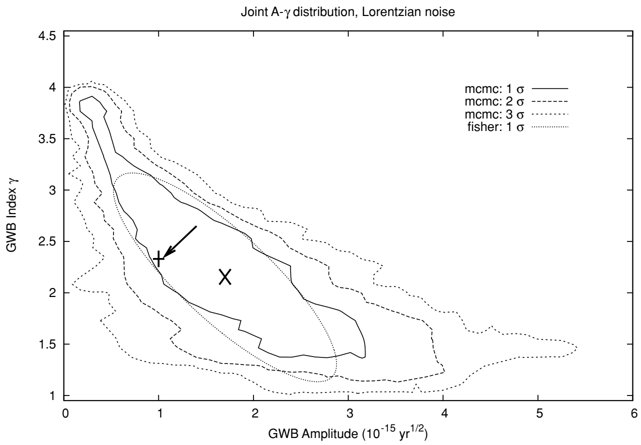

Figure 3. Same as in Fig. 2, but the mock data are generated and analysed using Lorentzian timing noise. Overall timing noise amplitude and characteristic frequency f_0 are taken to be 100 ns and 1 yr^{-1} for each pulsar.

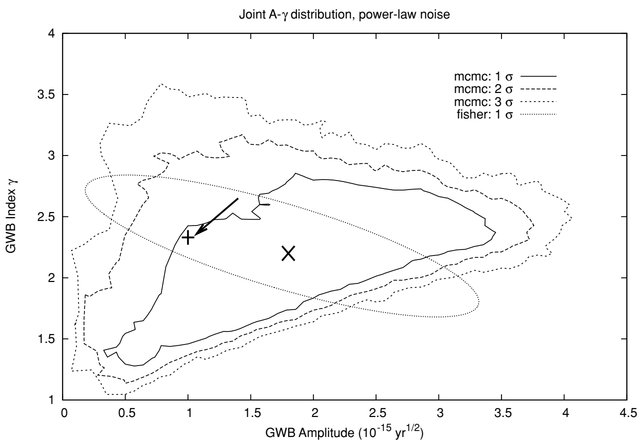

Figure 4. Same as in Fig. 2, but the mock data are generated and analysed using power-law timing noise. Overall timing noise amplitude and spectral index γ_i are taken to be 100 ns and 1.5 for each pulsar. For all γ_i, a prior distribution according to equation (27).

et al. 2009). We were able to generate data sets with exactly the same parameters as with our own algorithm, provided that the timing noise was white. We have confirmed that those data sets yield similar results when analysed with our algorithm.

An important point is that the spectral form of the timing noise has a large impact on the detectability of the GWB. For a red Lorentzian pulsar timing noise, there is far greater degeneracy between the spectral slope and amplitude in the TR data for the GWB than for white pulsar timing noise, and thus the overall S/N ratio is significantly reduced by the red component of the timing noise.

5.2 Multiple data sets, same input parameters

To estimate the robustness of our algorithm, we also perform a maximum likelihood search on many data sets.

(i) We generate a multitude of mock TR data for the same PTA configurations as in Section 5.1, with white timing noise.

(ii) For every one of these data sets, we calculate the maximum likelihood parameters using the conjugate directions search. The ensemble of maximum likelihood estimators for (A, γ) should be close to the true values used to generate the TRs.

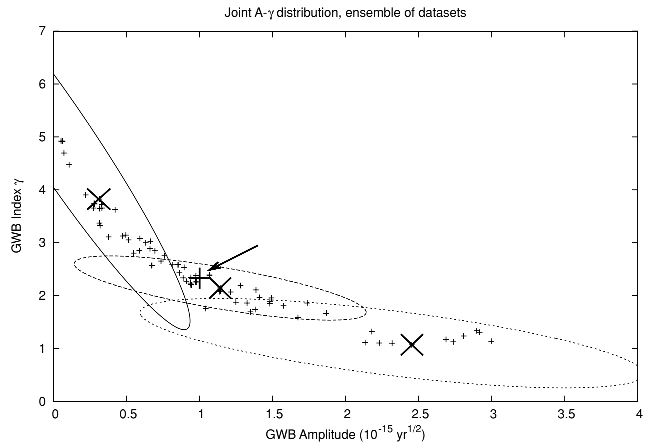

Figure 5. The maximum likelihood values for an ensemble of realizations of mock data sets, all with the same model parameters: 100 ns white noise, 20 pulsars and 100 data points per pulsar approximately evenly distributed over 5 years. The contours are confidence contours based on Fisher information matrix approximations of the likelihood function.

The results of maximum likelihood search on many data sets is demonstrated in Fig. 5. The points are the maximum likelihood values for individual data sets. It can be seen that the points are distributed in a shape similar, but not identical, to Fig. 2: some points are quite far off from the input parameters. In order to test the validity of the results, we calculate the Fisher information matrix at the maximum likelihood points and show the 1σ contour of the multidimensional Gaussian approximation based on the Fisher information matrix for three points. We see that the error contours do not exclude the true values at high confidence, even though the Fisher matrix does not yield a perfect representation of the error contours (the true posterior is not perfectly Gaussian), and we have a posteriori selected outliers for two of the three cases.

5.3 Parameter studies

To test the accuracy of the algorithm, and to provide suggestions for optimal PTA configurations, we conduct some parameter studies on simplified sets of mock TRs.

(a) We generate many sets of mock TRs for the simplified case of white pulsar timing noise spectra, all with the same white noise amplitude. The data sets are TRs for some number of millisecond pulsars which are positioned randomly in the sky. We parametrize the GWB by equation (32). We then generate many sets of TRs, varying several parameters [i.e. timing noise amplitude (assumed the same for all pulsars), duration of the experiment and number of pulsars].

(b) For each of the mock data sets, we approximate the likelihood function by a Gaussian in the GWB amplitude A, with all other parameters fixed to their real value. We use A as a free parameter since it represents the strength of the GWB, and therefore the accuracy of A is a measure of the detectability. All other parameters are fixed to keep the computational time low, but this does result in a higher S/N ratio than is obtainable with a full MCMC analysis.

(c) For this Gaussian approximation, we calculate the ratio μ/σ as an estimate for the S/N ratio, where μ is the value of A at which

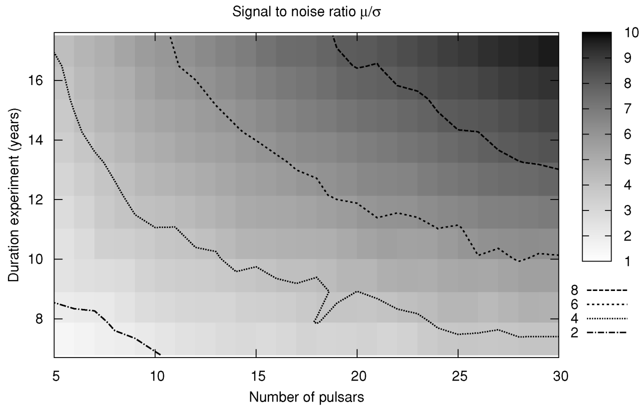

Figure 6. Density plot of the S/N ratio μ/σ for different realizations of TRs. We have assumed monthly observations of pulsars with white timing noise of 100 ns each. The GWB amplitude has been set to $10^{-15}\ \mathrm{yr}^{1/2}$.

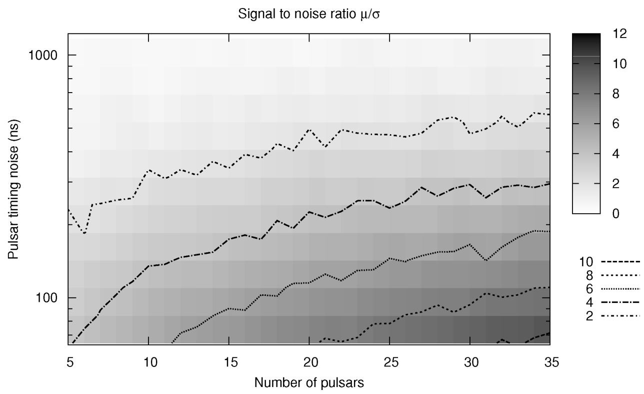

Figure 7. Density plot of the S/N ratio μ/σ for different realizations of TRs. We have assumed 100 data points per pulsars, approximately evenly distributed over a period of 7.5 years. The GWB amplitude has been set to $10^{-15}\ \mathrm{yr}^{1/2}$.

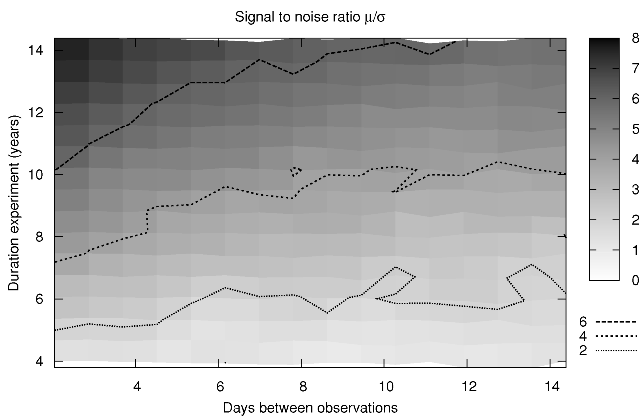

Figure 8. Density plot of the S/N ratio μ/σ for different realizations of TRs. We have used a constant GWB amplitude of $10^{-15}\ \mathrm{yr}^{1/2}$ and 20 pulsars.

the likelihood function maximizes and σ is the value of the standard deviation of the Gaussian approximation. Our results, represented as S/N ratio contour plots for pairs of the input parameters, can be seen in Figs 6–12.

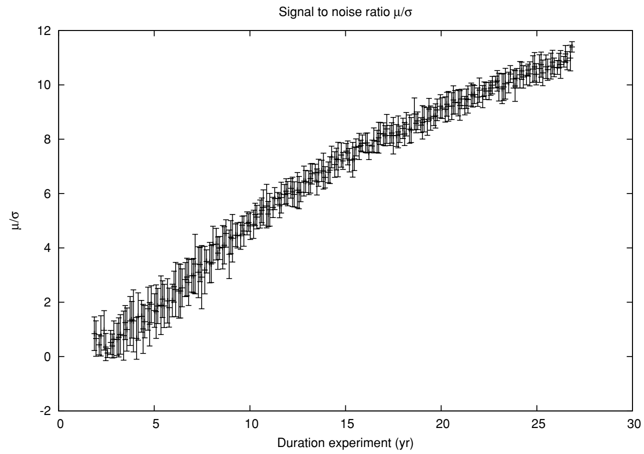

Figure 9. Density plot of the S/N ratio μ/σ for different realizations of TRs. We have used 20 pulsars with white pulsar timing noise levels of 100 ns each, with monthly observations. The GWB amplitude has been set to $10^{-15}\ \mathrm{yr}^{1/2}$. The points and error bars are the mean and standard deviation of 10 realizations.

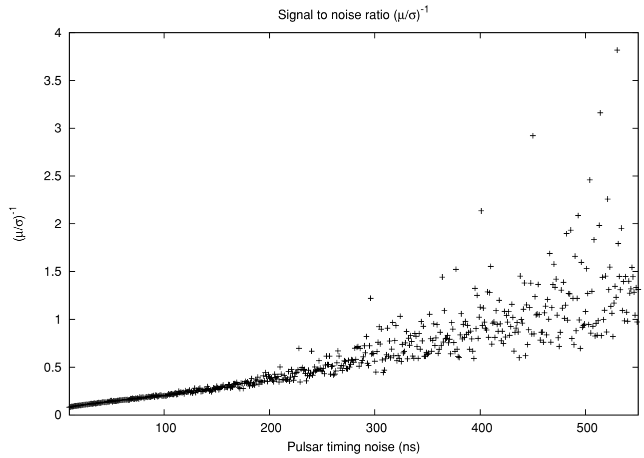

Figure 10. Plot of one over the S/N ratio $(\mu/\sigma)^{-1}$ with respect to the pulsar timing noise for an experiment of 5 years, 20 pulsars and monthly observations. The GWB amplitude has been set to $10^{-15}\ \mathrm{yr}^{1/2}$.

5.4 Comparison to other work

More than a decade ago, McHugh et al. (1996) used a Bayesian technique to produce upper limits on the GWB using pulsar timing.[1] We found the presentation of this work rather difficult to follow. None the less, it is clear that the analysis presented here is more general than that of McHugh et al.: we treat the whole pulsar array, and not just a single pulsar; we take into account the extreme redness of the noise and develop the formalism to treat the systematic errors like quadratic spin-down.

Simultaneously with our work, a paper by Anholm et al. (2008, hereafter A08) has appeared on the arxiv preprint service. Their approach was to construct a quadratic estimator (written explicitly in the frequency domain), which aims to be optimally sensitive to the GWB. This improves on the original non-quadratic estimator of J05. However, a number of issues important for the pulsar timing experiment remained unaddressed, the most important among

[1] We thank the anonymous referee for attracting our attention to this paper.

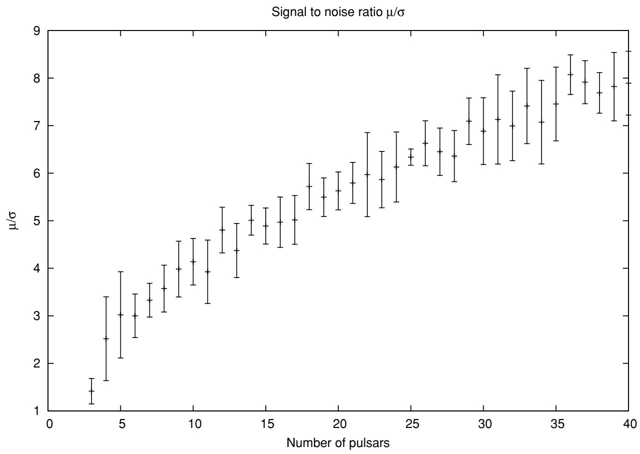

Figure 11. Plot of the S/N ratio μ/σ with respect to the number of observed pulsars. The white timing noise of each pulsar has been set to 100 ns and the observations were taking every 2 months for a period of 7.5 years. The GWB amplitude has been set to $10^{-15}\,\mathrm{yr}^{1/2}$. The points and error bars are the mean and standard deviation of 10 realizations.

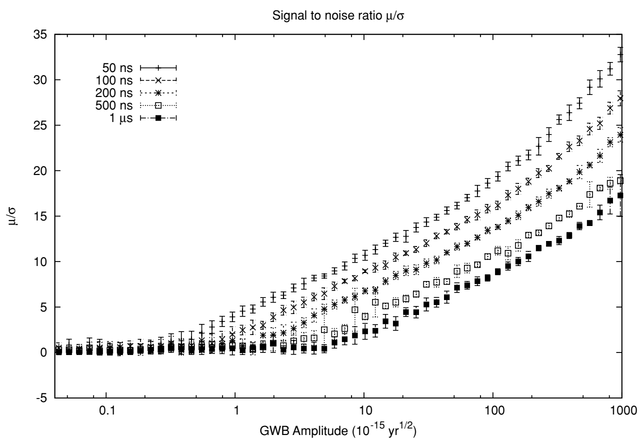

Figure 12. Several plots of the S/N ratio μ/σ with respect to the level of the GWB amplitude. The number of pulsars was set at 20, with biweekly observations for a period of 5 years. The pulsar noise levels were set at 50, 100, 200, 500, 1000 ns for the different plots. The points and error bars are the mean and standard deviation of 10 realizations.

them the extreme redness of the GWB and the need to subtract consistently the quadratic spin-down.

6 CONCLUSION

In this paper, we have introduced a practical Bayesian algorithm for measuring the GWB using PTAs. Several attractive features of the algorithm should make it useful to the PTA community:

(i) the ability to simultaneously measure the amplitude and slope of GWB,

(ii) its ability to deal with unevenly sampled data sets and

(iii) its ability to treat systematic contributions of known functional form. From the theoretical point of view, the algorithm is guaranteed to extract information optimally, provided that our parametrization of the timing noise is correct.

Test runs of our algorithm have shown that the experiments S/N ratio strongly decreases with the redness of the pulsar timing noise and strongly increases with the duration of the PTA experiment. We have also charted the S/N ratio dependence on the number of well-clocked pulsars and the level of their timing noise. These charts should be helpful in the design of the optimal strategy for future PTA observations.

ACKNOWLEDGMENTS

We thank Dan Stinebring, Dick Manchester, George Hobbs, Russell Edwards, Rick Jenet, Ben Stappers, Jason Hessels and Matthew Bailes for insightful discussions about the precision pulsar timing. RvH and YL thank ATNF for its annual hospitality. This research is supported by the Netherlands organisation for Scientific Research (NWO) through VIDI grant 639.042.607.

REFERENCES

Anholm M., Ballmer S., Creighton J. D. E., Price L. R., Siemens X., 2008, preprint (arXiv:0809.0701)
Begelman M. C., Blandford R. D., Rees M. J., 1980, Nat, 287, 307
Brent R., 1973, Algorithms for Minimization without Derivatives. Prentice-Hall, Englewood Cliffs, New Jersey
Britton M. C., 2000, ApJ, 532, 1240
Damour T., Vilenkin A., 2005, Phys. Rev. D, 71, 063510
Detweiler S., 1979, ApJ, 234, 1100
Efron B., 1979, Ann. Stat., 7, 1
Estabrook F., Wahlquist H., 1975, Gen. Relativ. Gravitation, 6, 439
Foster R., Backer D., 1990, ApJ, 361, 300
Grishchuk L. P., 2005, Phys. Usp., 48, 1235
Hellings R., Downs G., 1983, ApJ, 265, L39
Hobbs G. B., Edwards R. T., Manchester R. N., 2006, MNRAS, 369, 655
Hobbs G. et al., 2009, MNRAS, in press (arXiv:0901.0592)
Jaffe A., Backer D., 2003, ApJ, 583, 616
Jenet F., Hobbs G., Lee K., Manchester R., 2005, ApJ, 625, L123 (J05)
Jenet F. et al., 2006, ApJ, 653, 1571
Kaspi V. M., Taylor J. H., Ryba M. F., 1994, ApJ, 428, 713
Lommen A. N., 2001, PhD thesis, UC Berkeley
Maggiore M., 2000, Phys. Rev., 331, 283
Manchester R. N., 2006, Chin. J. Astron. Astrophys Suppl., 6, 139
McHugh M. P., Zalamansky G., Vernotte F., Lantz E., 1996, Phys. Rev. D, 54, 5993
Newman M., Barkema G., 1999, Monte Carlo Methods in Statistical Physics. Oxford Univ. Press, Oxford, p. 31
Phinney E. S., 2001, preprint (arXiv:0108028)
Roberts G., Gelman A., Gilks W., 1997, Ann. Appl. Prob., 7, 110
Sazhin M., 1978, SvA, 55, 65
Sesana A., Haardt F., Madau P., Volonteri M., 2005, Clas. Quantum Grav., 22, 363
Stappers B. W., Kramer M., Lyne A. G., D'Amico N., Jessner A., 2006, Chin. J. Astron. Astrophys. Suppl., 6, 298
Wyithe J., Loeb A., 2003, ApJ, 595, 614

APPENDIX A:

In this Appendix, we show explicitly how to perform marginalization over the nuisance parameters $\boldsymbol{\xi}$ in equation (16), rewritten here for convenience:

$$P(\sigma, \boldsymbol{\xi}|\delta\boldsymbol{x}) = \frac{M}{\sqrt{\det \mathbf{C}}} \exp\left[-\frac{1}{2}(\delta\boldsymbol{x} - \mathbf{F}\boldsymbol{\xi})\mathbf{C}^{-1}(\delta\boldsymbol{x} - \mathbf{F}\boldsymbol{\xi})\right] \times P_0(\sigma, \boldsymbol{\xi}), \tag{A1}$$

From here on, we will assume that P_0 is independent of $\boldsymbol{\xi}$ (a flat prior). All values are therefore equally likely for all elements of $\boldsymbol{\xi}$ prior to the observations. This assumption is also implicitly made in the frequentist approach when fitting for these kinds of parameters as is done in popular pulsar timing packages. We now perform the marginalization:

$$P(\sigma|\delta\boldsymbol{x}) = \int P(\sigma, \boldsymbol{\xi}|\delta\boldsymbol{x})\, \mathrm{d}^m\xi, \tag{A2}$$

where m is the dimensionality of $\boldsymbol{\xi}$. The idea now is to rewrite the exponent E of equation (A1) in such a way that we can perform a Gaussian integral with respect to $\boldsymbol{\xi}$ (we have to get rid of the $\mathbf{F}$ in front of $\boldsymbol{\xi}$). Therefore, we will expand E and complete the square with respect to ξ:

$$\begin{aligned} E &= (\delta\boldsymbol{x} - \mathbf{F}\boldsymbol{\xi})^T\, \mathbf{C}^{-1}\, (\delta\boldsymbol{x} - \mathbf{F}\boldsymbol{\xi}) \\ &= \delta\boldsymbol{x}^T\mathbf{C}^{-1}\delta\boldsymbol{x} - 2\boldsymbol{\xi}^T\mathbf{F}^T\mathbf{C}^{-1}\delta\boldsymbol{x} + \boldsymbol{\xi}^T\mathbf{F}^T\mathbf{C}^{-1}\mathbf{F}\boldsymbol{\xi} \\ &= \delta\boldsymbol{x}^T\mathbf{C}^{-1}\delta\boldsymbol{x} + (\boldsymbol{\xi} - \boldsymbol{\chi})^T\, \mathbf{F}^T\mathbf{C}^{-1}\mathbf{F}\, (\boldsymbol{\xi} - \boldsymbol{\chi}) \\ &\quad - \boldsymbol{\chi}^T\mathbf{F}^T\mathbf{C}^{-1}\mathbf{F}\boldsymbol{\chi}, \end{aligned} \tag{A3}$$

where we have used the substitution:

$$\boldsymbol{\chi} = (\mathbf{F}^T\mathbf{C}^{-1}\mathbf{F})^{-1}\mathbf{F}^T\mathbf{C}^{-1}\delta\boldsymbol{x}. \tag{A4}$$

Using this, we can write the $\boldsymbol{\xi}$ dependent part of the integral of equation (A2) as a multidimensional Gaussian integral:

$$\begin{aligned} I &= \int \exp\left[\frac{-1}{2}\, (\boldsymbol{\xi} - \boldsymbol{\chi})^T\, \mathbf{F}^T\mathbf{C}^{-1}\mathbf{F}\, (\boldsymbol{\xi} - \boldsymbol{\chi})\right] \mathrm{d}^m\xi \\ &= (2\pi)^m \det\left(\mathbf{F}^T\mathbf{C}^{-1}\mathbf{F}\right)^{-1}. \end{aligned} \tag{A5}$$

From this it follows that

$$\begin{aligned} P(\sigma|\delta\boldsymbol{x}) = &\frac{M'}{\sqrt{\det(\mathbf{C})\det(\mathbf{F}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{F})}} \\ &\times \exp\left[-\frac{1}{2}\delta\boldsymbol{x}\mathbf{C}'\delta\boldsymbol{x}\right], \end{aligned} \tag{A6}$$

where we have absorbed all constant terms in the normalization constant M' and where we have used

$$\mathbf{C}' = \mathbf{C}^{-1} - \mathbf{C}^{-1}\mathbf{F}(\mathbf{F}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{F})^{-1}\mathbf{F}^{\mathrm{T}}\mathbf{C}^{-1}. \tag{A7}$$

This paper has been typeset from a TEX/LATEX file prepared by the author.