

# Including the Pulsar-Term in Continuous Gravitational-wave Searches using Pulsar Timing Arrays: A Blessing and a Curse

(Dated: February 13, 2014)

The last several years has seen a growing effort to develop robust and powerful data-analysis techniques for the purpose of detection and characterisation of gravitational-waves (GWs) using pulsar timing arrays (PTAs). When a GW passes the Earth-pulsar line of sight, it causes a perturbation to the intervening space-time metric which may leave a measurable imprint on the time-of-arrival (TOA) of radio pulses from regularly observed millisecond pulsars. By timing an array of these pulsars, we create a Galactic-scale GW detector, sensitive between  $\sim 1 - 100$  nHz. The most likely signal at lower frequencies will be a stochastic background formed from the incoherent superposition of signals from inspiraling SuperMassive Black-Hole Binaries (SMBHBs). However at higher frequencies the stochasticity of the signal begins to break down, and we may be able to resolve individual SMBHB sources through detection of their continuous GW emission. Here we describe several new techniques (including the  $\mathcal{M}_p$  statistic) which collapse the dimensionality of Bayesian continuous-wave searches, mitigating the increase of dimensionality with the size of the pulsar array, and making the analysis tractable with powerful evidence-evaluation packages like MULTINEST. We find acceptable agreement of our techniques with the parameter-estimation and Bayes' factor evaluation performed with the full signal template, and conclude that these techniques make excellent first-cut tools for detection and characterisation of continuous GW signals with PTAs.

PACS numbers:

## I. INTRODUCTION

The rotational-stability and exquisite timing precision of many Galactic millisecond pulsars (MSPs) affords us the unique opportunity of establishing a network of “clocks” in space [1]. With robust models of pulse-arrival times, any observed deviations have the potential to encode the signature of transiting extra-galactic gravitational waves [2–5]. Effectively acting as single arms of an interferometer, precision timing of MSPs has the potential to open up the  $10^{-9}$ – $10^{-7}$  Hz band of GW frequencies, a regime hitherto inaccessible to ground- and space-based detectors, and of particular interest since it uniquely includes the early inspiral regime of supermassive black-hole (SMBH) binaries [6].

## II. THE SIGNAL MODEL

Repeated observation of radio signals from a pulsar allows us to construct a timing-model describing all deterministic contributions to the observed pulse times-of-arrival (TOA). We define pulsar timing-residuals as the difference between the observed TOAs and the TOAs of a fitted deterministic timing-model. These residuals encode all unmodelled phenomena which influence the TOAs, such as various noise processes (detector noise, intrinsic spin-noise, DM, etc.), and extragalactic GWs which perturb the intervening Earth-pulsar space-time metric. Here, we provide a brief review of the model for the timing-residuals induced by a circular, non-spinning SMBHB.

The GW is a perturbation to the flat space-time metric, which, in the transverse traceless (TT) gauge, can be written as,

$$h_{ab}(t, \hat{\Omega}) = e_{ab}^+(\hat{\Omega})h_+(t, \hat{\Omega}) + e_{ab}^\times(\hat{\Omega})h_\times(t, \hat{\Omega}) \quad (1)$$

where  $\hat{\Omega}$  is the direction of propagation of the GW, and  $h_+$ ,  $h_\times$  and  $e_{ab}^A$  ( $A = +, \times$ ) are the GW polarisation amplitudes and basis-tensors, respectively. Using the notation of Wahlquist [7], the independent polarisation basis-tensors can be written as,

$$e_{ab}^+(\hat{\Omega}) = m_a m_b - n_a n_b, \quad e_{ab}^\times(\hat{\Omega}) = m_a n_b + n_a m_b \quad (2)$$

where,

$$\begin{aligned} \hat{\Omega} &= -(\sin \theta \cos \phi) \hat{x} - (\sin \theta \sin \phi) \hat{y} - (\cos \theta) \hat{z} \\ \hat{m} &= -(\sin \phi) \hat{x} + (\cos \phi) \hat{y} \\ \hat{n} &= -(\cos \theta \cos \phi) \hat{x} - (\cos \theta \sin \phi) \hat{y} + \sin \theta \hat{z} \end{aligned} \quad (3)$$

where  $(\theta, \phi) = (\pi/2 - \text{DEC}, \text{RA})$  denotes the sky-location of the source. The GW-induced pulsar timing-residuals are written as,

$$s(t, \hat{\Omega}) = F^+(\hat{\Omega})\Delta s_+(t) + F^\times(\hat{\Omega})\Delta s_\times(t) \quad (4)$$

where  $\Delta s_A(t) = s_A(t_p) - s_A(t_e)$ , and  $t_{p,e}$  denote the times at which the GW rolls past the pulsar and the Earth, respectively. The functions  $F^A(\hat{\Omega})$  are known as the antenna pattern functions, encoding the geometrical sensitivity of a particular pulsar to a propagating GW, and are defined as,

$$F^+(\hat{\Omega}) = \frac{1}{2} \frac{(\hat{m} \cdot \hat{p})^2 - (\hat{n} \cdot \hat{p})^2}{1 + \hat{\Omega} \cdot \hat{p}} \quad (5)$$

$$F^\times(\hat{\Omega}) = \frac{(\hat{m} \cdot \hat{p})(\hat{n} \cdot \hat{p})}{1 + \hat{\Omega} \cdot \hat{p}}$$

where  $\hat{p}$  is a unit vector pointing from the Earth to the pulsar. With simple geometrical arguments, we can write  $t_p = t_e - L(1 + \hat{\Omega} \cdot \hat{p})$ , where  $L$  is the distance to the pulsar<sup>1</sup>.

With these definitions, we can write the GW-induced pulsar timing-residuals for a circular, non-spinning SMBHB as [7–9],

$$s_+(t) = \frac{\mathcal{M}^{5/3}}{D_L \omega(t)^{1/3}} \left[ -\sin[2(\Phi(t) - \phi_n)] (1 + \cos^2 \iota) \cos 2\psi - 2 \cos[2(\Phi(t) - \phi_n)] \cos \iota \sin 2\psi \right] \quad (6)$$

$$s_\times(t) = \frac{\mathcal{M}^{5/3}}{D_L \omega(t)^{1/3}} \left[ -\sin[2(\Phi(t) - \phi_n)] (1 + \cos^2 \iota) \sin 2\psi + 2 \cos[2(\Phi(t) - \phi_n)] \cos \iota \cos 2\psi \right]$$

where  $\psi$  is the GW polarisation angle,  $\iota$  is the binary orbital-inclination angle,  $\phi_n$  is the orbital phase at the line of nodes,  $\mathcal{M}$  is the chirp mass of the binary (defined in terms of the component masses of the binary as  $\mathcal{M} = (m_1 m_2)^{3/5} / (m_1 + m_2)^{1/5}$ ), and  $D_L$  is the luminosity distance to the source. We collect  $\Phi_0$  and  $\phi_n$  into one constant *initial phase* variable,  $\phi_0 = \phi_n - \Phi_0$ . Equations (4) and (6) can now be used together to define the induced pulsar-timing residuals for a non-evolving or evolving SMBHB.

### A. Non-evolving signal template

Consider the low-frequency regime, where evolution of the source frequency is small, such that the frequencies of the GW when it passes the pulsar and the Earth are approximately the same. We can include the pulsar-term in our single-source template by modelling the signal as the sum of two sinusoids of different phases. The signal template in the  $\alpha^{\text{th}}$  pulsar, where  $\omega_0$  is the orbital angular-frequency of an inspiralling SMBHB GW-source, is,

$$s_\alpha = \sum_{i=1}^2 a_{i\alpha}(\zeta, \iota, \psi, \phi_0, \phi_\alpha, \theta, \phi) A_\alpha^i(t, \omega_0) \quad (7)$$

where,

$$a_{1\alpha} = [(F_\alpha^+ a_1 + F_\alpha^\times a_3) (1 - \cos \phi_\alpha) - (F_\alpha^+ a_2 + F_\alpha^\times a_4) \sin \phi_\alpha]$$

$$a_{2\alpha} = [(F_\alpha^+ a_2 + F_\alpha^\times a_4) (1 - \cos \phi_\alpha) + (F_\alpha^+ a_1 + F_\alpha^\times a_3) \sin \phi_\alpha] \quad (8)$$

and,

---

<sup>1</sup> Using units such that  $G = c = 1$ .

$$\begin{aligned}
a_1 &= \zeta \left[ (1 + \cos^2 \iota) \cos 2\phi_0 \cos 2\psi + 2 \cos \iota \sin 2\phi_0 \sin 2\psi \right] \\
a_2 &= -\zeta \left[ (1 + \cos^2 \iota) \sin 2\phi_0 \cos 2\psi - 2 \cos \iota \cos 2\phi_0 \sin 2\psi \right] \\
a_3 &= \zeta \left[ (1 + \cos^2 \iota) \cos 2\phi_0 \sin 2\psi - 2 \cos \iota \sin 2\phi_0 \cos 2\psi \right] \\
a_4 &= -\zeta \left[ (1 + \cos^2 \iota) \sin 2\phi_0 \sin 2\psi + 2 \cos \iota \cos 2\phi_0 \cos 2\psi \right].
\end{aligned} \tag{9}$$

In the above equations,  $\phi_\alpha = 2\omega_0 L_\alpha (1 + \hat{\Omega} \cdot \hat{p})$  and  $\zeta = \mathcal{M}^{5/3}/D_L$ . The signal basis-functions are defined as,

$$A_\alpha^1 = \frac{1}{\omega_0^{1/3}} \sin(2\omega_0 t), \quad A_\alpha^2 = \frac{1}{\omega_0^{1/3}} \cos(2\omega_0 t). \tag{10}$$

Now, the log-likelihood ratio of the data given some signal parameters is,

$$\ln \Lambda = \sum_{\alpha=1}^{N_p} (r_\alpha | s_\alpha) - \frac{1}{2} (s_\alpha | s_\alpha) \tag{11}$$

where the inner product is defined by  $(x|y) = x^T G (G C G)^{-1} G^T y$ ,  $C$  is a noise-matrix,  $G$  is a matrix which performs an analytic marginalisation over timing-model parameters [10], and  $r_\alpha$  is a vector of timing-residuals in the  $\alpha^{\text{th}}$  pulsar.

We now write  $(s_\alpha | s_\alpha)$  and  $(r_\alpha | s_\alpha)$  explicitly in terms of the pulsar-phase parameters,  $\phi_\alpha$ . Let  $p_\alpha = (F_\alpha^+ a_1 + F_\alpha^\times a_3)$  and  $q_\alpha = (F_\alpha^+ a_2 + F_\alpha^\times a_4)$ . Therefore,

$$\begin{aligned}
(s_\alpha | s_\alpha) &= \left[ p_\alpha^2 (1 - \cos \phi_\alpha)^2 + q_\alpha^2 \sin^2 \phi_\alpha - 2 p_\alpha q_\alpha (1 - \cos \phi_\alpha) \sin \phi_\alpha \right] (A_\alpha^1 | A_\alpha^1) \\
&\quad + \left[ q_\alpha^2 (1 - \cos \phi_\alpha)^2 + p_\alpha^2 \sin^2 \phi_\alpha + 2 p_\alpha q_\alpha (1 - \cos \phi_\alpha) \sin \phi_\alpha \right] (A_\alpha^2 | A_\alpha^2) \\
&\quad + 2 \left[ p_\alpha q_\alpha (1 - \cos \phi_\alpha)^2 + p_\alpha^2 (1 - \cos \phi_\alpha) \sin \phi_\alpha - q_\alpha^2 (1 - \cos \phi_\alpha) \sin \phi_\alpha - p_\alpha q_\alpha \sin^2 \phi_\alpha \right] (A_\alpha^1 | A_\alpha^2) \\
(r_\alpha | s_\alpha) &= [p_\alpha (r_\alpha | A_\alpha^1) + q_\alpha (r_\alpha | A_\alpha^2)] - [p_\alpha (r_\alpha | A_\alpha^1) + q_\alpha (r_\alpha | A_\alpha^2)] \cos \phi_\alpha - [q_\alpha (r_\alpha | A_\alpha^1) - p_\alpha (r_\alpha | A_\alpha^2)] \sin \phi_\alpha.
\end{aligned} \tag{12}$$

These expressions can be plugged into Eq. (11), giving us an expression for the log-likelihood ratio written explicitly in terms of the pulsar-phase variables,  $\phi_\alpha$ . In a parameter-estimation search or an evaluation of the Bayes-factor, conventional techniques would require a search over  $7 + N_{\text{psr}}$  dimensions. For large arrays or expensive likelihood evaluations this can be a costly exercise, necessitating multi-threading linear-algebra operations to accelerate the likelihood evaluations, or multi-core machines to perform efficient parallel-tempering for the evaluation of Bayes factors. One should also note that the popular and effective Bayesian inference tool MULTINEST has little use in these kinds of high-dimensional problems (even in constant efficiency mode), as the set of live-points used in the nested sampling algorithm very slowly accumulates the last few units of log-evidence<sup>2</sup>.

### 1. Techniques for maximisation/marginalisation over $\phi_\alpha$

By explicitly exposing  $\phi_\alpha$  in the likelihood-ratio, we have developed several alternative approaches designed to approximate the likelihood-ratio maximised/marginalised over these pulsar-phase variables. Firstly, as noted in Ellis et al. [9], one can avoid the formalism of the  $\mathcal{F}_p$  statistic (which maximises the likelihood-ratio over  $2N_{\text{psr}}$  “amplitude” parameters [ $a_{i\alpha}$  in Eq. (7)] despite there being only  $7 + N_{\text{psr}}$  independent parameters). The  $\mathcal{F}_p$  statistic is most likely an appropriate approximation for small pulsar arrays, however as we include more pulsars the disparity between the dimensionality of the parameter-space assumed by the  $\mathcal{F}_p$  statistic and the true *physical* parameter-space grows larger.

To avoid maximising over these nuisance “amplitude” parameters, Ellis et al. [9] note that it is possible to analytically maximise over the physical  $\phi_\alpha$  parameters. This requires solving a quartic equation in  $x = \cos \phi_\alpha$  which is guaranteed

---

<sup>2</sup> Certain alternative approaches to this have been proposed for MULTINEST, such as the use of importance nested sampling in constant-efficiency mode [11], or employing a trained neural network [12] to accelerate the final stages of sampling.

to have at least one unique solution, although whether that solution satisfies the requirement  $-1 \leq x \leq 1$  would have to be ascertained on the fly. We can of course, avoid this completely by *numerically* maximising over the pulsar-phase parameters. This is **Technique #1**, and constitutes a more appropriate maximisation than  $\mathcal{F}_p$ . Nevertheless, we are still left with the problem of searching over the remaining 7-dimensional parameter space; this is a much more tractable problem and can be handled with many off-the-shelf MCMC or nested-sampling algorithms. In this case, we should not be surprised if any bias is observed in the posterior distributions of the final 7 parameters, since we are after all *maximising* over  $N_{\text{psr}}$  other parameters. Bayesian model-selection is inappropriate in this case, where maximisation over the parameter space has been employed.

The second option is to avoid maximising entirely, and to *numerically* marginalise the likelihood-ratio over the pulsar-phase parameters, such that we actually sample the marginalised likelihood-ratio in our MCMC or nested-sampling algorithms. In particular, if we can do this without increasing the likelihood evaluation time significantly, then the collapse of the dimensionality makes this problem tractable with MULTINEST. There are many benefits to this; for example, MULTINEST is an excellent tool for sampling multimodal distributions, it has inbuilt parallelisation, and in low-dimensionality provides an efficient means to evaluate the Bayesian evidence. Hence, the numerical marginalisation of the non-evolving signal-model over pulsar-phase parameters is our **Technique #2**.

Finally, we note that, if we have sufficiently many wave cycles during the observation time of the pulsars in our array, it is possible to maximise over the pulsar-phase parameters analytically without the need to solve a quartic. More interestingly, it is also possible to analytically marginalise over the pulsar-phase parameters. The noise behaviour of real pulsars and the GW frequencies to which we are most sensitive will likely prohibit us from making the assumptions required to analytically maximise/marginalise in this fashion. However, we provide the derivation and a brief discussion in the Appendix.

## B. Evolving signal model

We have developed a similar technique with an evolving-signal template; this takes into account evolution of the SMBHB orbital frequency over the Earth-pulsar light-travel time, but still assumes evolution during the actual pulsar observation time is negligible.

For this evolving-signal template, we define,

$$\begin{aligned} A_\alpha^1 &= \frac{1}{\omega_0^{1/3}} \sin(2\omega_0 t), & A_\alpha^2 &= \frac{1}{\omega_0^{1/3}} \cos(2\omega_0 t) \\ B_\alpha^1 &= \frac{1}{\omega_{p,\alpha}^{1/3}} \sin(2\omega_{p,\alpha} t), & B_\alpha^2 &= \frac{1}{\omega_{p,\alpha}^{1/3}} \cos(2\omega_{p,\alpha} t). \end{aligned} \quad (13)$$

where  $\omega_{p,\alpha} = \omega_0 - \dot{\omega}_0 L_\alpha (1 + \hat{\Omega} \cdot \hat{p}_\alpha)$ ,  $\dot{\omega}_0 = (96/5)\mathcal{M}^{5/3}\omega_0^{11/3}$ , and  $L_\alpha$  is the distance to the  $\alpha^{\text{th}}$  pulsar.

Expressing the log-likelihood ratio explicitly in terms of the pulsar-phase parameters ( $\phi_\alpha = 2\omega_0 L_\alpha (1 + \hat{\Omega} \cdot \hat{p}_\alpha)$ ) gives the following,

$$\begin{aligned} \ln \Lambda &= \sum_{\alpha=1}^{N_p} [-p_\alpha(r_\alpha|B_\alpha^1) - q_\alpha(r_\alpha|B_\alpha^2) + p_\alpha^2(A_\alpha^1|B_\alpha^1) + q_\alpha^2(A_\alpha^2|B_\alpha^2) + p_\alpha q_\alpha \{(A_\alpha^1|B_\alpha^2) + (A_\alpha^2|B_\alpha^1)\}] \cos \phi_\alpha \\ &+ [-q_\alpha(r_\alpha|B_\alpha^1) + p_\alpha(r_\alpha|B_\alpha^2) + q_\alpha^2(A_\alpha^2|B_\alpha^1) - p_\alpha^2(A_\alpha^1|B_\alpha^2) + p_\alpha q_\alpha \{(A_\alpha^1|B_\alpha^1) - (A_\alpha^2|B_\alpha^2)\}] \sin \phi_\alpha \\ &- \frac{1}{2} [p_\alpha^2(B_\alpha^1|B_\alpha^1) + q_\alpha^2(B_\alpha^2|B_\alpha^2) + 2p_\alpha q_\alpha (B_\alpha^1|B_\alpha^2)] \cos^2 \phi_\alpha \\ &- \frac{1}{2} [q_\alpha^2(B_\alpha^1|B_\alpha^1) + p_\alpha^2(B_\alpha^2|B_\alpha^2) - 2p_\alpha q_\alpha (B_\alpha^1|B_\alpha^2)] \sin^2 \phi_\alpha \\ &- [p_\alpha q_\alpha ((B_\alpha^1|B_\alpha^1) - (B_\alpha^2|B_\alpha^2)) - p_\alpha^2(B_\alpha^1|B_\alpha^2) + q_\alpha^2(B_\alpha^1|B_\alpha^2)] \sin \phi_\alpha \cos \phi_\alpha \\ &+ p_\alpha(r_\alpha|A_\alpha^1) + q_\alpha(r_\alpha|A_\alpha^2) - \frac{1}{2} p_\alpha^2(A_\alpha^1|A_\alpha^1) - \frac{1}{2} q_\alpha^2(A_\alpha^2|A_\alpha^2) - p_\alpha q_\alpha (A_\alpha^1|A_\alpha^2). \end{aligned} \quad (14)$$

There are two ways to proceed with this signal template, but both involve numerical marginalisation over the pulsar-phase parameters. In **Technique #3**, in order to compute  $\omega_{p,\alpha}$  we fix  $L_\alpha$  to its catalogued value, while in **Technique #4** we internally average over the prior distribution of  $L_\alpha$  by drawing the distance used to compute  $\omega_{p,\alpha}$  from a Gaussian centred on the catalogued value with standard-deviation given by the catalogued error-bars.

TABLE I: Pulsar distances taken from Verbiest et al. [14] if available, or otherwise from the ATNF catalogue [15].

Pulsar	White-noise RMS / ns	Time-span / yr	Pulsar distances / kpc
J0030+0451	792	12.7	$0.28 \pm 0.1$
J0437-4715	69	14.8	$0.156 \pm 0.001$
J1640+2224	410	14.9	$1.19 \pm 0.238$
J1713+0747	136	18.3	$1.05 \pm 0.06$
J1744-1134	366	16.9	$0.42 \pm 0.02$
J1857+0943	402	14.9	$0.9 \pm 0.2$
J1909-3744	100	9.0	$1.26 \pm 0.03$
J1939+2134	141	16.3	$5.0 \pm 2.0$
J2317+1439	412	14.9	$1.89 \pm 0.38$

Even though the pulsar-phase has a dependence on the pulsar-distance, including the distance in parameter estimation can produce practical difficulties, as a change in the distance may have a relatively small effect on the pulsar frequency,  $\omega_{p,\alpha}$ , but can have a huge impact on the phase coherence [13]. By using our approximation of marginalising over the pulsar-phase, and drawing  $L_\alpha$  from within its prior to calculate  $\omega_{p,\alpha}$ , we side-step this problem.

### III. RESULTS

While a full analysis of these techniques in all conceivable situations is beyond the scope of this paper, we rigorously test what we expect to be the most promising new technique. **Technique #4** (which from now we denote as the  $\mathcal{M}_p$  statistic) is subjected to a program of systematic injection and recovery of simulated signals, using the PALSIMULATION code which is part of the PAL package [16] being developed as a unifying suite of tools for pulsar timing analysis. Being a similar technique, the performance of **Technique #3** closely follows that of **Technique #4**, while we expect no systematic bias from **Technique #2** other than that which is introduced by analysing an evolving signal with a non-evolving formalism.

The datasets we generate are of the following configurations;

- **Type I:** 36 pulsars, 5 years of observations, 2 week cadence, 100 ns RMS white-noise per pulsar,  $L_{\text{psr}} = 1 \pm 0.1$  kpc  $\forall$  pulsars; effectively IPTA MDC Open1 format
- **Type II:** 9 pulsars, variable observation time-span, average 2 week cadence, realistic white-noise,  $L_{\text{psr}}$  equal to catalogued values
- **Type III:** 9 pulsars, variable observation time-span, average 2 week cadence, realistic white-noise,  $L_{\text{psr}}$  drawn from Gaussian distribution (mean=catalogued-value, standard-deviation=catalogued-error)

The observation time-spans, white-noise RMS values, and distances for the 9 pulsars in Type II and Type III datasets are shown in Table I.

#### A. Bayesian model selection

An important issue in GW data analysis is the strength of evidence for the alleged detection of a signal. Within the context of Bayesian inference, this is quantified by the posterior odds ratio,

$$\frac{p(\mathcal{H}_2|\vec{D})}{p(\mathcal{H}_1|\vec{D})} = \frac{p(\vec{D}|\mathcal{H}_2)p(\mathcal{H}_2)}{p(\vec{D}|\mathcal{H}_1)p(\mathcal{H}_1)} = \frac{\mathcal{Z}_2 \times p(\mathcal{H}_2)}{\mathcal{Z}_1 \times p(\mathcal{H}_1)}. \quad (15)$$

where  $p(\mathcal{H}|\vec{D})$  is the probability of a model given data,  $p(\vec{D}|\mathcal{H})$  is the probability of data given a model (more commonly known as the evidence,  $\mathcal{Z}$ ), and  $p(\mathcal{H})$  is the prior probability for a given model. The prior probability ratio for the two competing models can often be set to one, such that the posterior odds ratio is equated to the Bayes' factor,  $\mathcal{B} = \mathcal{Z}_2/\mathcal{Z}_1$ .

The Bayesian evidence,  $\mathcal{Z}$ , necessarily involves an integral over the entire parameter space,

$$\mathcal{Z} = \int \mathcal{L}(\vec{\mu})\pi(\vec{\mu})d^N\mu, \quad (16)$$

and as such computing it can be a costly exercise. However nested-sampling [17], as implemented in the Bayesian inference package MULTINEST [11, 18, 19], provides an efficient means to evaluate  $\mathcal{Z}$ , with the added benefit of also providing posterior distribution samples for parameter estimation.

We now note that plugging the *likelihood-ratio* into Eq. (16) rather than simply the likelihood gives us a direct evaluation of the Bayes' factor for a signal-versus-null model-comparison. Bearing in mind that we assume a fixed-noise search with no stochastic-search parameters such that  $\det[2\pi GCG]$  is constant, we see,

$$\begin{aligned} \mathcal{B} = \frac{\mathcal{Z}_{\text{signal}}}{\mathcal{Z}_{\text{null}}} &= \frac{\int \left(1/\sqrt{\det[2\pi GCG]}\right) \exp[-(r-s(\vec{\mu})|r-s(\vec{\mu}))/2]\pi(\vec{\mu})d^N\mu}{\int \left(1/\sqrt{\det[2\pi GCG]}\right) \exp[-(r|r)/2]\pi(\vec{\mu})d^N\mu} \\ &= \frac{\int \exp[(r|s(\vec{\mu})) - (1/2)(s(\vec{\mu})|s(\vec{\mu}))]\pi(\vec{\mu})d^N\mu}{\int \pi(\vec{\mu})d^N\mu} \\ &= \int \Lambda(\vec{\mu})\pi(\vec{\mu})d^N\mu \end{aligned} \quad (17)$$

We evaluate the veracity of the Bayes' factors returned by these pulsar-phase marginalisation techniques by injecting signals into various noise realisations at various SNRs. We compare the returned Bayes' factors with those returned by employing parallel-tempering with the full signal model. The PTA datasets are again of Type I, II and III.

The signal we inject matches that explored in Ellis [13], which is at the sky-location of the Fornax cluster. Recent work has shown that there may be potential single GW source “hot spots” in the Virgo, Fornax and Coma clusters [20]. Regardless, we are only interested in sensible parameters to form an injected signal. These parameters are;  $\{\mathcal{M}, D_L, f_0, \phi, \cos\theta, \cos\iota, \psi, \phi_0\} = \{7 \times 10^8 M_\odot, -, 10^{-8}\text{Hz}, 0.95, -0.56, 0, 1.26, 0.99\}$ , where the luminosity distance  $D_L$  is scaled to suit the desired SNR.

Another important aspect is our choice of prior on  $\mathcal{M}$ ,  $D_L$  and  $f_0$ . We employ log-uniform priors on these variables, but also apply a cut on the characteristic-strain induced by the binary such that the signal does not become large enough to see “by eye” in the residuals. Defining  $h_0 = 4\sqrt{2/5}\omega_0^{2/3}\zeta$ , we apply the cut  $h_0 < (f_{\text{gw}}/10^{-8}\text{Hz})^{3/2}h_{0,c}$ , where  $h_{0,c} = 10^{-13}$ . We use numerical integration to re-normalise the prior to incorporate this cut in parameter space, which is essential for the correct computation of Bayes' factors. In practice, we find that the prior re-normalisation leads to a change in log odds-ratio which is less than 0.1.

The comparison between an evaluation of the posterior odds ratio performed by the full thermodynamic integration (solid lines) and the  $\mathcal{M}_p$  statistic (dashed lines) for Type I/II/III datasets of various injected SNR is shown in Fig. 1, where we see excellent agreement for a variety of different noise realisations. We find that the speed of the numerical-marginalisation techniques depend on the SNR of the injection, where for low to moderate SNR ( $\sim 0-2$ ) the evidence and parameter-estimation stages of MultiNest completed within only a few minutes of wall-time on 48 computational cores. The highest SNR injections (SNR = 10) required longer, but still finished within  $\sim 45$  minutes of wall-time on 48 cores.

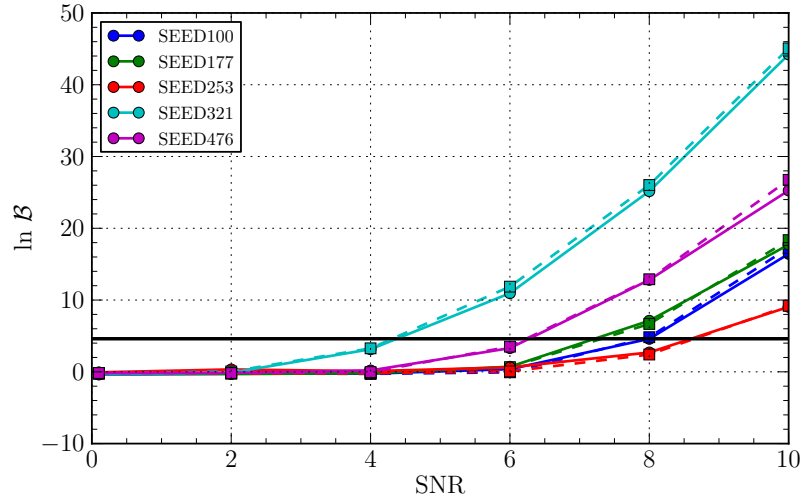
Analysing these datasets using the numerical phase marginalisation with a non-evolving signal-model (**Technique #2**), we find that the mismatch between the model and the evolving-signal injections leads to Bayes' factors which can be significantly below the optimal evolving-model values. This is illustrated for a single noise realisation in Figure 2. We will revisit this in the next section.

## B. Bayesian parameter estimation

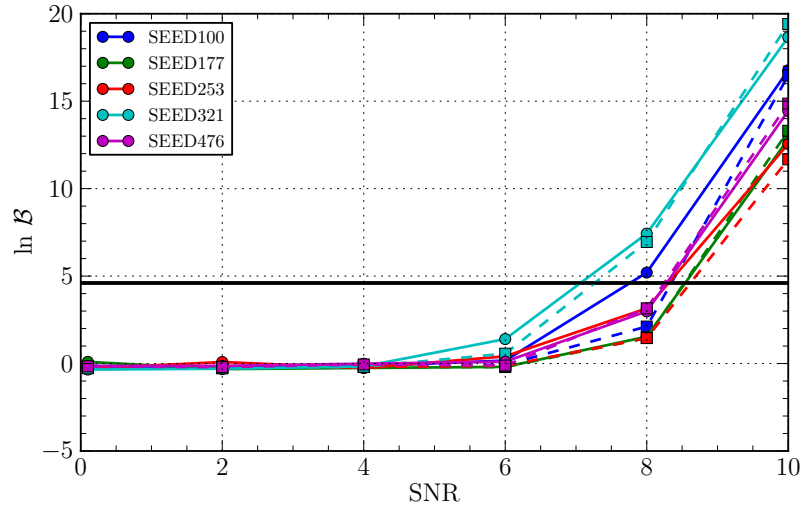
To ascertain whether numerical marginalisation techniques introduce any systematic bias in parameter recovery, we inject SNR = 8 signals into various white-noise realisations. The injected binary orbital frequency is chosen to be  $10^{-8}$  Hz such that the GW frequency lies close to the peak sensitivity of an array of pulsars observed over a period of  $\gtrsim 5$  years. We choose injected chirp masses of  $7 \times 10^8 M_\odot$  and  $1.8 \times 10^8 M_\odot$  in order to model a strongly evolving (over the Earth-pulsar light-travel time), and weakly evolving binary respectively, where the lower mass injection will have a frequency derivative which is  $\sim 10\%$  of the higher mass.

These evolving and weakly-evolving binaries are injected into 100 different noise-realizations of Type II datasets. This type of dataset is used because we want the characteristics of the PTA to remain fixed, such that the injected binary's luminosity distance,  $D_L$ , (which is scaled to accommodate the desired SNR) is constant over each realisation. The remaining binary parameters are injected with the following values into each dataset;  $\{\phi = 1, \cos\theta = 0.48, \cos\iota = 0.88, \psi = 0.5, \phi_0 = 0.5\}$ .

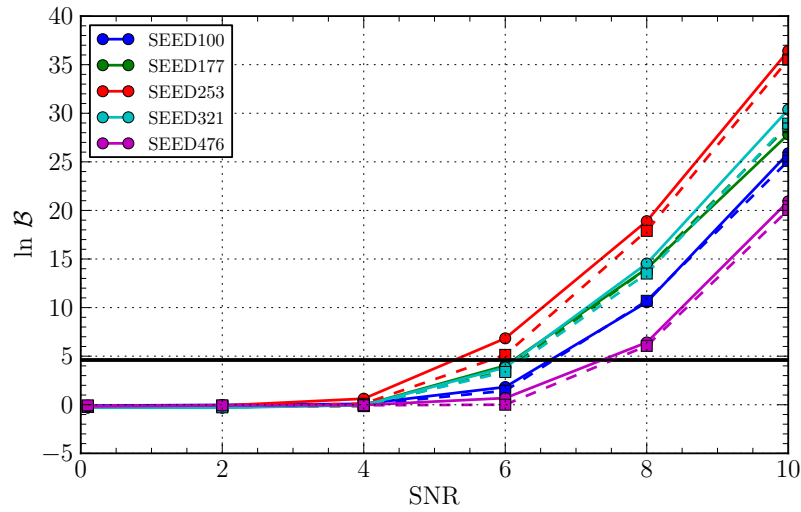
We present results for the case of the  $\mathcal{M}_p$  statistic, which should apply regardless of whether the binary is evolving or not. We again note that no bias would be expected within **Technique #2**, which numerically marginalises over



(a)



(b)



(c)

FIG. 1: A comparison of the computed posterior odds-ratios ( $\ln \mathcal{B}$ ) evaluated using thermodynamic integration of the full signal model (solid lines), and the technique of numerically marginalising the pulsar-phase parameters while sampling from the pulsar-distance prior ( $\mathcal{M}_p$  statistic; dashed lines). Different SNR signals are injected into a variety of realisations of Type I/II/III datasets. The agreement found between the two methods is excellent.

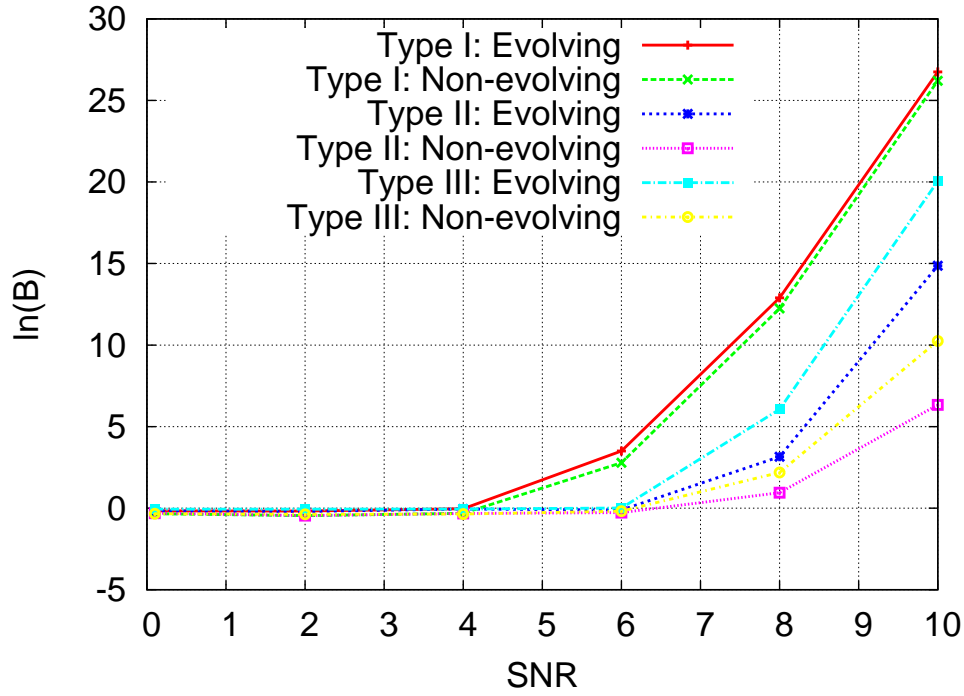


FIG. 2: For a given realisation of noise, we repeat the analysis of Type I/II/III datasets with a non-evolving signal-model. We see that the mismatch between the assumption of a non-evolving signal and the reality of an evolving-binary injection leads to Bayes' factors which can be significantly below the optimal evolving-model values.

the pulsar-phase variables in the non-evolving formalism. The only bias expected here derives from the inherent limitations of applying an inappropriate non-evolving model to a possibly evolving signal.

Our method of testing for systematic bias in the use of the  $\mathcal{M}_p$  statistic is an extension of a method used in Ellis et al. [21] to validate the accuracy of a first-order likelihood approximation in a stochastic background search. As discussed there, the benchmark of internal consistency is when, in  $x\%$  of realisations, the set of injected parameters lies within the inner  $x\%$  of the marginalised posterior distribution. The inner high-probability region is defined as,

$$\int_W p(\vec{\theta}) d^N \theta = a \quad (18)$$

$$W = \{\theta^1, \theta^2, \dots, \theta^N \in \mathbb{R} : p(\vec{\theta}) > L_a\}$$

where  $L_a > 0$  is some value unique to each  $a$  corresponding to a curve of equal probability in the  $N$  dimensional parameter space.

To find all points satisfying  $p(\vec{\theta}) > L_a$  we rank the recovered posterior samples in terms of decreasing posterior weight, then integrate over all samples until we reach the desired credible interval. For each realisation, we can then define two sets of points; the set of points inside the high-probability region (HPR)  $\mathcal{S}_a$ , and the complementary set  $\bar{\mathcal{S}}_a$ .

We now extend the dimensionality of the definitions of the  $\chi^2$  variables in [21] to give a measure of the distance of the posterior samples in each set from the true injected parameters,



$$\begin{aligned}
\chi_a(\vec{\theta}_i)^2 &= \left( \frac{\log_{10}(\mathcal{M}_i) - \log_{10}(\mathcal{M}_{\text{true}})}{\log_{10}(\mathcal{M}_{\text{true}})} \right)^2 + \left( \frac{\log_{10}(D_{L,i}) - \log_{10}(D_{L,\text{true}})}{\log_{10}(D_{L,\text{true}})} \right)^2 + \left( \frac{\phi_i - \phi_{\text{true}}}{\phi_{\text{true}}} \right)^2 \\
&\quad + \left( \frac{\cos \theta_i - \cos \theta_{\text{true}}}{\cos \theta_{\text{true}}} \right)^2 + \left( \frac{\cos \iota_i - \cos \iota_{\text{true}}}{\cos \iota_{\text{true}}} \right)^2 + \left( \frac{\psi_i - \psi_{\text{true}}}{\psi_{\text{true}}} \right)^2 + \left( \frac{\phi_{0,i} - \phi_{0,\text{true}}}{\phi_{0,\text{true}}} \right)^2 \\
\bar{\chi}_a(\vec{\theta}_j)^2 &= \left( \frac{\log_{10}(\mathcal{M}_j) - \log_{10}(\mathcal{M}_{\text{true}})}{\log_{10}(\mathcal{M}_{\text{true}})} \right)^2 + \left( \frac{\log_{10}(D_{L,j}) - \log_{10}(D_{L,\text{true}})}{\log_{10}(D_{L,\text{true}})} \right)^2 + \left( \frac{\phi_j - \phi_{\text{true}}}{\phi_{\text{true}}} \right)^2 \\
&\quad + \left( \frac{\cos \theta_j - \cos \theta_{\text{true}}}{\cos \theta_{\text{true}}} \right)^2 + \left( \frac{\cos \iota_j - \cos \iota_{\text{true}}}{\cos \iota_{\text{true}}} \right)^2 + \left( \frac{\psi_j - \psi_{\text{true}}}{\psi_{\text{true}}} \right)^2 + \left( \frac{\phi_{0,j} - \phi_{0,\text{true}}}{\phi_{0,\text{true}}} \right)^2
\end{aligned} \tag{19}$$

where  $\vec{\theta}_i, \vec{\theta}_j$  are elements of the sets  $\mathcal{S}_a$  and  $\mathcal{S}_{\bar{a}}$  respectively. Finally, we define the empirical distribution function (EDF) as,

$$F_k(a) = \frac{1}{k} \sum_{n=1}^k \Theta(\min \chi_a^2 - \min \chi_a^2) \tag{20}$$

where  $\Theta(x)$  is the Heaviside step-function. The term inside the sum indicates an instance of a noise-realisation where the injected values are “closer” (in the  $\chi^2$  sense) to one of the elements of the HPR than to any element of the complementary set.

The results of such an analysis is shown in Fig. 3 for the evolving and weakly-evolving binary injections. The line of internal consistency is shown as a thick, black diagonal line. We see that this technique does indeed present bias, with a worst-case sag of  $\sim 0.25$ . However, the EDF does not give an insight into how this bias manifests itself in the parameter space.

We now show the distribution of maximum-a-posterior values over all 100 noise-realizations, with the injected signal parameters also indicated. Figure 4 indicates that, while the  $\mathcal{M}_p$  statistic may fail the formal EDF test, in practical terms it quite comfortably recovers the true parameters of the injected signal. Even when our catalogue of pulsar distances is offset from the true values by an amount consistent with their error-bars, we can see from Fig. 5 that we still accurately recover the injected signal parameters. This is also true when the source is weakly evolving. Figure 6 shows the same analysis for a weakly evolving injection, where, despite some offset of  $(\mathcal{M}, D_L)$  from the injected values which may be consistent with the scatter of maximum-a-posteriori values, we again achieve an accurate recovery of the signal parameters.

A further test we can perform is to assess the performance of the Bayesian posterior odds-ratio as a detection classifier within this technique. We do so by producing a receiver operator characteristic (ROC) plot, illustrating the fraction of true positive detections versus false positive detections as we vary the detection threshold. We inject various SNR signals into 100 different noise-realizations, recovering the evidence in each case. The injected binary parameters are the same as the evolving case above. Therefore, we see from Fig. 7 that the posterior odds-ratio becomes a virtually perfect detection classifier at an SNR of 6. Although we can not draw truly general conclusion from this, the aim of this exercise is to show that these numerical marginalisation techniques are rapid and accurate enough to allow these kinds of detailed statistical tests within a Bayesian context. The question of what is required for an unambiguous Bayesian-claim of GW detection can be rigorously assessed by employing these techniques.

Finally, we assess the importance of using an evolving versus non-evolving signal-model when establishing detection criteria. For evolving and weakly-evolving sources, we inject SNR=8 signals into 100 different noise-realizations. We analyse each dataset using both the numerical phase marginalisation in the evolving-model ( $\mathcal{M}_p$  statistic) and the non-evolving model, recovering the evidences in each case. The results are shown in Fig. 8, where we see that the evolving signal-model is more general, capturing the behaviour of the gravitational-waveform even when the signal is non-evolving, and giving a Bayes’ factor which is comparable to the value returned by the non-evolving analysis. However, as seen in the previous section, the non-evolving signal-model performs poorly when applied to evolving source injections, and returning a Bayes’ factor which can be significantly lower than the optimal evolving-model values.

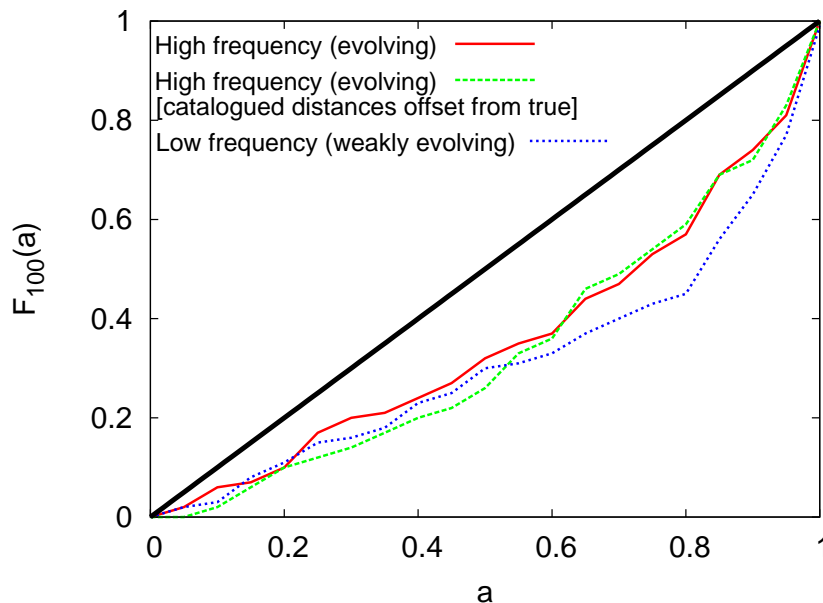


FIG. 3: The fraction of injections which are “closer” in the chi-squared sense (see text) to the set of points lying inside credible-interval,  $a$ , is plotted against the credible-interval. The line of zero-bias is shown as a thick, black-line, while the results of an analysis of 100 realisations of evolving/non-evolving Type II datasets using numerical-marginalisation (the  $\mathcal{M}_p$  statistic) are shown as solid-red and dashed-blue. The dashed-green line shows the result for when we offset our catalogue of distances from their true values by an amount consistent with their error-bars. While some bias is present, this plot does not indicate how that manifests in the physical parameter-space.

## IV. CONCLUSIONS

### Appendix

#### 1. Analytic marginalisation/maximisation over $\phi_\alpha$ (non-evolving)

Assuming we have sufficiently many wave cycles during the observation time-span, then we can use the following assumptions for the signal basis-function overlaps:  $(A_1|A_2) = (A_2|A_1) \simeq 0$ , and  $(A_1|A_1) \simeq (A_2|A_2) \simeq \mathcal{N}(\omega_0)$ . In practice, the ratio of the cross-terms of the basis-function overlaps to the diagonal terms may not be small enough to permit these approximations to be used. For example, Fig. 9 shows the ratio  $(A_1|A_2)/(A_1|A_1)$  for one of the pulsars in the IPTA MDC Open1 dataset, and for a real NANOGrav J0613-0200 dataset [22]. The ratio diminishes at higher frequencies, and for the mock dataset gets to  $\lesssim 10^{-2}$  at the highest frequencies we are sensitive to. However, for a real pulsar dataset the ratio stays around  $10^{-1}$  even at the highest frequencies. Furthermore, the GW frequencies to which we are most sensitive are  $\sim \text{a few } \times 10^{-8}$ , diminishing as we move to the higher frequencies required for these approximations to hold.

Nevertheless, these analytic expressions may have some value as rapid first-pass tools, and we provide the derivations below.

#### a. Marginalising

Given the overlap approximations, and the non-evolving signal-template in Sec. IIA,

$$\begin{aligned} (s_\alpha|s_\alpha) &\simeq [a_{1\alpha}a_{1\alpha} + a_{2\alpha}a_{2\alpha}]\mathcal{N}(\omega_0), \\ &\simeq 2\mathcal{N}(\omega_0) (p_\alpha^2 + q_\alpha^2) (1 - \cos \phi_\alpha) \end{aligned} \quad (21)$$

Thus,

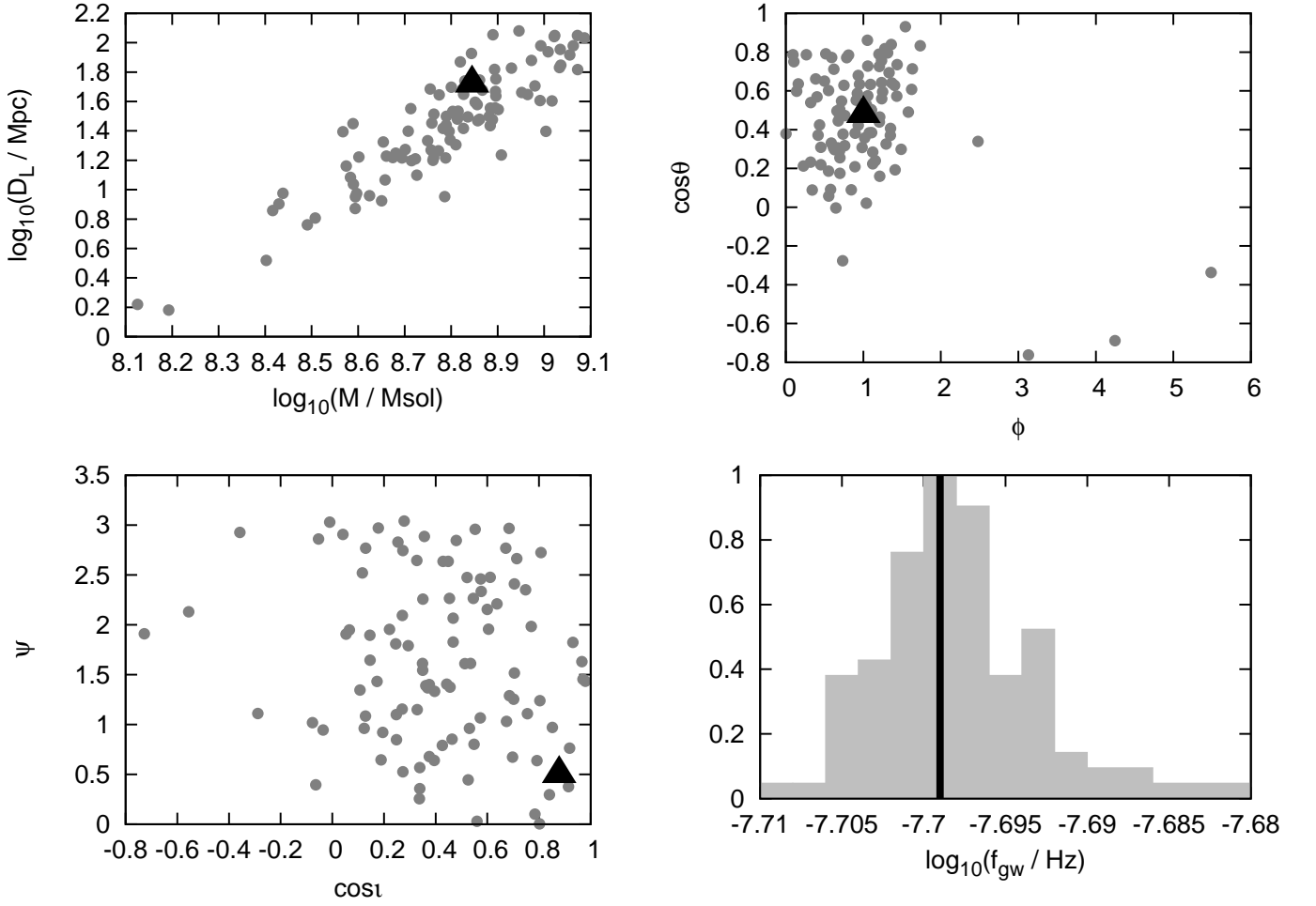


FIG. 4: We show the distribution of maximum-a-posteriori values (filled grey circles) from an analysis of 100 realisations of an evolving signal injected into a Type II dataset, and analysed with the  $\mathcal{M}_p$  statistic. As can be seen, in the parameters of interest ( $\mathcal{M}, D_L, f_{\text{gw}}, \phi, \cos \theta, \cos i$ ) this technique recovers the injected values (filled black triangles) quite comfortably.

$$\begin{aligned}
(r_\alpha | s_\alpha) - \frac{1}{2}(s_\alpha | s_\alpha) &\simeq [p_\alpha(r_\alpha | A_\alpha^1) + q_\alpha(r_\alpha | A_\alpha^2) - (p_\alpha^2 + q_\alpha^2) \mathcal{N}(\omega_0)] \\
&\quad - [p_\alpha(r_\alpha | A_\alpha^1) + q_\alpha(r_\alpha | A_\alpha^2) - (p_\alpha^2 + q_\alpha^2) \mathcal{N}(\omega_0)] \cos \phi_\alpha \\
&\quad - [q_\alpha(r_\alpha | A_\alpha^1) - p_\alpha(r_\alpha | A_\alpha^2)] \sin \phi_\alpha \\
&\simeq -X_\alpha + X_\alpha \cos \phi_\alpha + Y_\alpha \sin \phi_\alpha
\end{aligned} \tag{22}$$

Hence, marginalising the likelihood-ratio over each pulsar-phase parameter, assuming flat-priors, gives,

$$\begin{aligned}
\int \Lambda d^{N_p} \phi &\propto \left( \frac{1}{2\pi} \right)^{N_p} \prod_{\alpha=1}^{N_p} \int_0^{2\pi} \exp[(r_\alpha | s_\alpha) - \frac{1}{2}(s_\alpha | s_\alpha)] d\phi_\alpha \\
&\propto \left( \frac{1}{2\pi} \right)^{N_p} \exp \left( \sum_{\alpha=1}^{N_p} -X_\alpha \right) \prod_{\alpha=1}^{N_p} \int_0^{2\pi} \exp(X_\alpha \cos \phi_\alpha + Y_\alpha \sin \phi_\alpha) d\phi_\alpha \\
&\propto \prod_{\alpha=1}^{N_p} \exp(-X_\alpha) I_0 \left( \sqrt{X_\alpha^2 + Y_\alpha^2} \right).
\end{aligned} \tag{23}$$

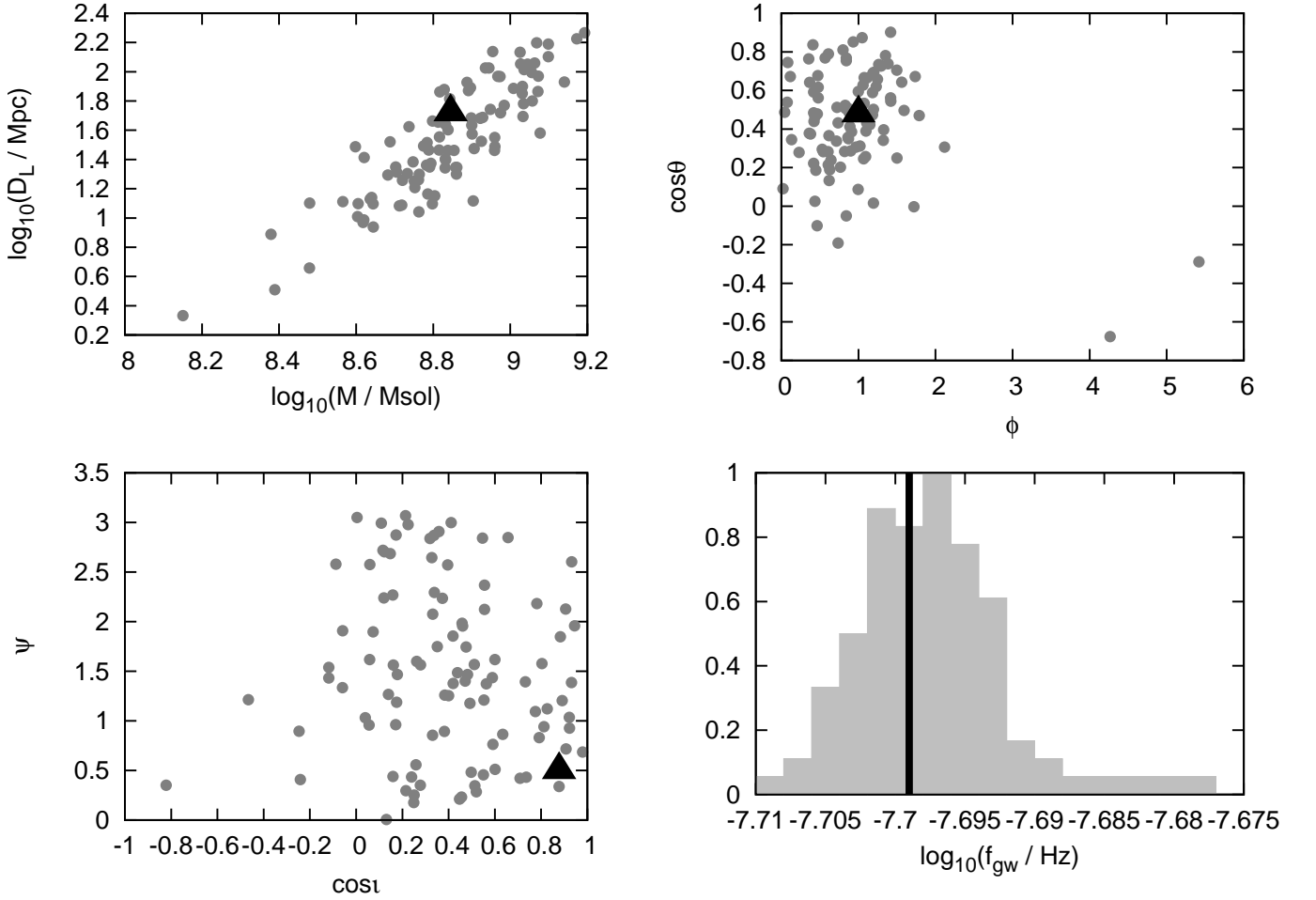


FIG. 5: We show the distribution of maximum-a-posteriori values (filled grey circles) from an analysis of the same 100 realisations of an evolving signal as in Fig. 4, but analysed assuming a catalogue of pulsar distances which are offset from their true values by an amount consistent with error bars. We see in this case that the  $\mathcal{M}_p$  statistic still comfortably recovers the injected parameters, despite having an offset catalogue of pulsar distances.

where  $I_0$  is a modified Bessel function of the first kind. Note that this technique of analytic marginalisation of nuisance phase parameters has previously been used in different contexts [23, 24]. Finally we have the PML (Phase Marginalised Likelihood) statistic,

$$\ln \Lambda_{\text{marg}} \propto \sum_{\alpha=1}^{N_p} \left\{ -X_{\alpha} + \ln \left[ I_0 \left( \sqrt{X_{\alpha}^2 + Y_{\alpha}^2} \right) \right] \right\} \quad (24)$$

If we have a bright source (with a large amplitude), such that the argument of the modified Bessel function is large, then directly computing  $I_0(x)$  can be very difficult. However, we can use a large argument expansion of the modified Bessel function to aid this calculation,

$$\ln [I_0(x)] \sim x - \frac{1}{2} \ln(2\pi x) + \ln \left( 1 + \frac{1}{8x} + \frac{9}{128x^2} + \frac{225}{3072x^3} + \frac{11025}{98304x^4} \cdots \right). \quad (25)$$

#### b. Maximising

Going back to the original  $\ln \Lambda$ , it is possible to maximise the likelihood-ratio over the pulsar-phase parameters. As indicated in Ellis et al. [9], the solution to the maximum-likelihood value of  $\phi_{\alpha}$  requires evaluating a quartic.

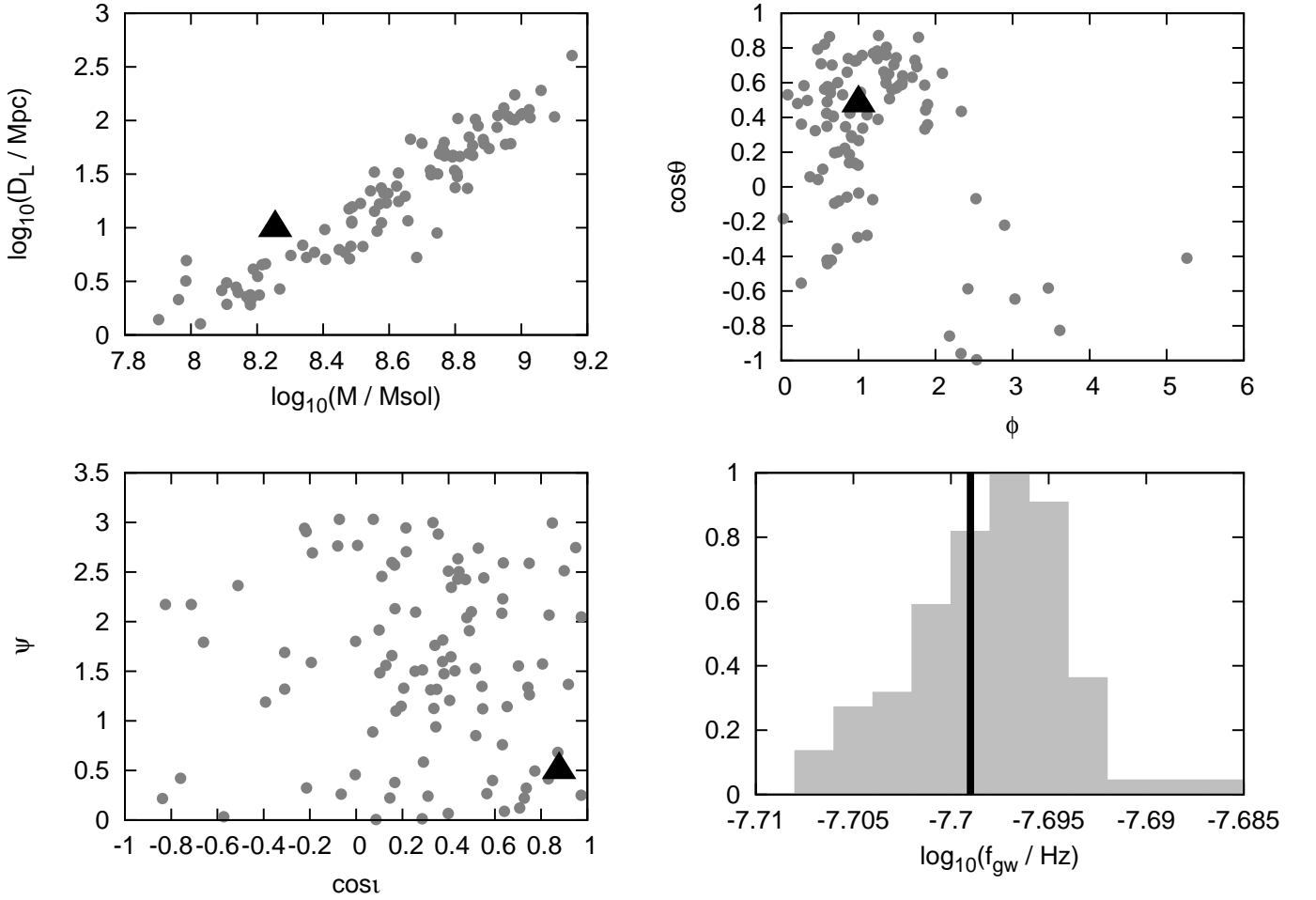


FIG. 6: We show the distribution of maximum-a-posteriori values (filled grey circles) from an analysis of 100 realisations of a weakly evolving signal injected into a Type II dataset, and analysed with the  $\mathcal{M}_p$  statistic. The injected values of  $(\mathcal{M}, D_L)$  appear to be offset from the distribution of maximum-a-posteriori values which may be consistent with the scatter of those values.

However, if we use the  $B_{1,2}$  overlap approximations, as in the previous section, then the solution is more simple. Maximising,

$$\frac{\partial \ln \Lambda}{\partial \phi_\beta} \simeq -X_\beta \sin \phi_\beta + Y_\beta \cos \phi_\beta = 0 \quad (26)$$

Hence,

$$\tan \phi_\beta = \frac{Y_\beta}{X_\beta} \quad (27)$$

such that we can define the log-likelihood ratio maximised over all  $\phi_\alpha$ , which we call the  $\mathcal{T}_p$ -statistic,

$$\mathcal{T}_p = \sum_{\alpha=1}^{N_p} \left[ -X_\alpha + \sqrt{X_\alpha^2 + Y_\alpha^2} \right] \quad (28)$$

We may be able to go further, and to maximise over other parameters. Regardless, we have a rather compact form for the log-likelihood ratio maximised over all the pulsar-phase parameters. The remaining 7-D single-source parameter space can easily be explored using MCMC.

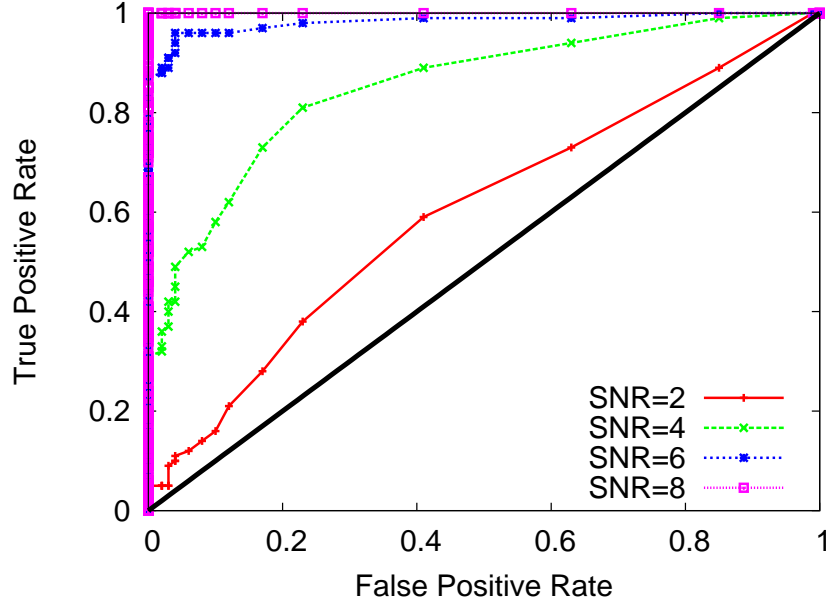


FIG. 7: We inject an evolving signal into 100 realisations of Type II datasets at various SNRs (including SNR=0), recovering the posterior-odds ratio via the  $\mathcal{M}_p$  statistic in each case. Setting the threshold of detection at varying values of the posterior odds ratio, we compute the fraction of realisations which are classified as false-positive and true-positive detections. We see that for this binary, and using this technique, the posterior odds ratio is an almost perfect classifier at SNR=6. With these numerical marginalisation techniques, the run-time is fast enough to permit detailed analysis of detection requirements within a Bayesian context.

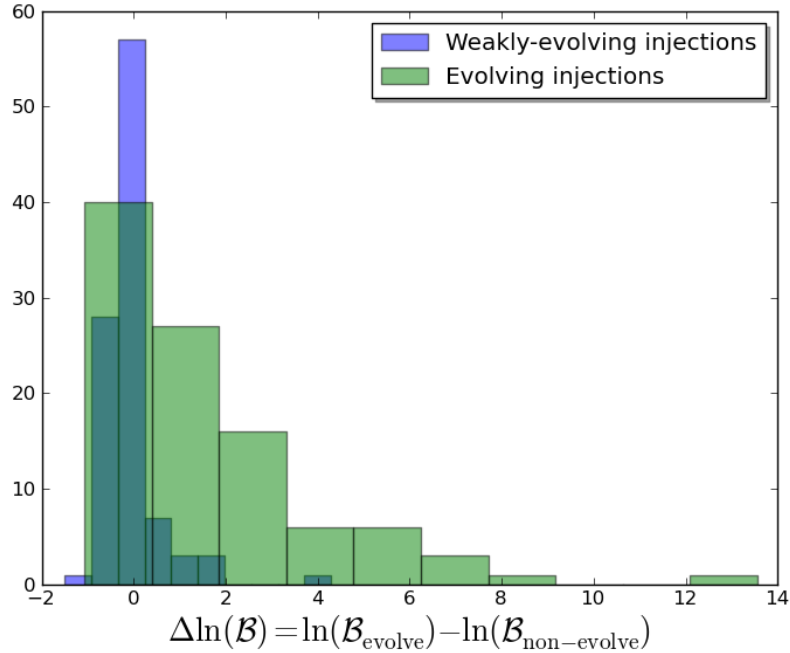


FIG. 8: Evolving and weakly-evolving signals are injected into 100 different noise-realizations with an SNR of 8. We analyse all datasets using both the non-evolving and evolving signal-models (with numerical phase marginalisation), recovering the evidence in each case. We find that the non-evolving and evolving models perform similarly when the signal is weakly-evolving, but the non-evolving model performs significantly worse than the evolving model whenever the signal is truly evolving, as expected.

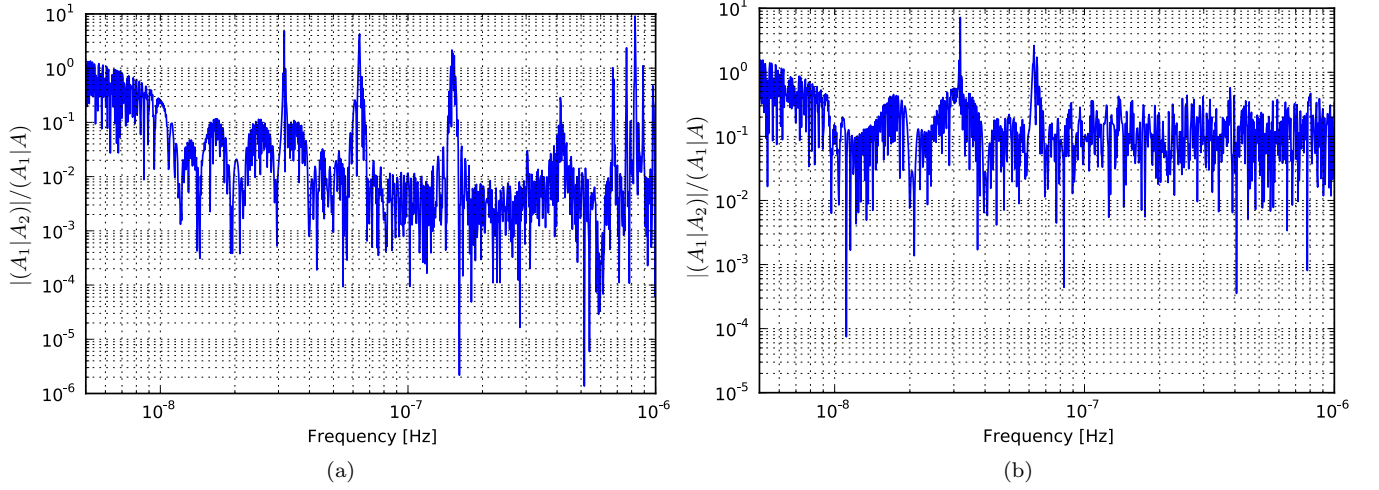


FIG. 9: The ratio of the basis-function overlaps in the cross-terms,  $(A_1|A_2)/(A_1|A_1)$ , is shown for (a) an IPTA MDC Open1 pulsar; 100 ns RMS white-noise, 2 week cadence; (b) a real NANOGrav dataset for J0613-0200 [22], where the noise is also fairly white.

Note that if we use the large argument expansion of the modified Bessel function to approximate the PML we get,

$$\ln \Lambda_{\text{marg}} \propto \sum_{\alpha=1}^{N_p} \left\{ -X_{\alpha} + \sqrt{X_{\alpha}^2 + Y_{\alpha}^2} - \frac{1}{2} \ln \left( 2\pi \sqrt{X_{\alpha}^2 + Y_{\alpha}^2} \right) \right\}. \quad (29)$$

For sufficiently large arguments,  $\sqrt{X_{\alpha}^2 + Y_{\alpha}^2}$  increases faster than  $\ln \left( 2\pi \sqrt{X_{\alpha}^2 + Y_{\alpha}^2} \right)$ . Hence, in the limit of a very strong signal the PML statistic is proportional to the  $\mathcal{T}_p$  statistic,

$$\begin{aligned} \ln \Lambda_{\text{marg}} &\propto \sum_{\alpha=1}^{N_p} \left\{ -X_{\alpha} + \sqrt{X_{\alpha}^2 + Y_{\alpha}^2} \right\} \\ &\propto \mathcal{T}_p. \end{aligned} \quad (30)$$

## 2. $\mathcal{B}_p$ statistic (analytic marginalisation over $a_{i\alpha}$ ; non-evolving)

Rather than analytically maximise over the amplitude parameters,  $a_{i\alpha}$ , to produce the  $\mathcal{F}_p$  statistic, if we assume uniform priors on these parameters then it is trivial to analytically marginalise over them,

$$\begin{aligned} \ln \Lambda &= \sum_{\alpha=1}^{N_p} (r_{\alpha}|s_{\alpha}) - \frac{1}{2} (s_{\alpha}|s_{\alpha}) \\ &= \sum_{\alpha=1}^{N_p} a_{i\alpha} (r_{\alpha}|A_{\alpha}^i) - \frac{1}{2} a_{i\alpha} a_{j\alpha} (A_{\alpha}^i|A_{\alpha}^j) \\ &= \sum_{\alpha=1}^{N_p} a_{i\alpha} N_{\alpha}^i - \frac{1}{2} a_{i\alpha} a_{j\alpha} M_{\alpha}^{ij} \\ &= \sum_{\alpha=1}^{N_p} a_{\alpha} N_{\alpha} - \frac{1}{2} a_{\alpha}^T M_{\alpha} A_{\alpha} \end{aligned} \quad (31)$$

Now, completing the square,

$$\ln \Lambda = -\frac{1}{2} \sum_{\alpha=1}^{N_p} \left[ (a_\alpha - M_\alpha^{-1} N_\alpha)^T M_\alpha (a_\alpha - M_\alpha^{-1} N_\alpha) - N_\alpha^T (M_\alpha^{-1})^T N_\alpha \right]. \quad (32)$$

Finally, integrating over the amplitude parameters with uniform priors,

$$\begin{aligned} \mathcal{B}_p &= C \exp \left( \sum_{\alpha=1}^{N_p} \frac{N_\alpha^T (M_\alpha^{-1})^T N_\alpha}{2} \right) \prod_{\alpha=1}^{N_p} \left[ \det (M_\alpha^{-1}) \times (2\pi)^2 \right]^{1/2} \\ &= C (2\pi)^{N_p} \exp(\mathcal{F}_p) \prod_{\alpha=1}^{N_p} (\det M_\alpha)^{-1/2} \end{aligned} \quad (33)$$

where  $C$  is a constant denoting the prior range of  $a_{i\alpha}$ . The form of the  $\mathcal{B}_p$  has been previously arrived at in the context of LIGO data analysis [25], but may have limited use in PTA analysis.

### Acknowledgments

S.R.T is supported by the STFC. J.R.G is supported by the Royal Society. This work was performed using the Darwin Supercomputer of the University of Cambridge High Performance Computing Service (<http://www.hpc.cam.ac.uk/>), provided by Dell Inc. using Strategic Research Infrastructure Funding from the Higher Education Funding Council for England.

- 
- [1] R. S. Foster and D. C. Backer, *ApJ* **361**, 300 (1990).
  - [2] F. B. Estabrook and H. D. Wahlquist, *General Relativity and Gravitation* **6**, 439 (1975).
  - [3] W. L. Burke, *ApJ* **196**, 329 (1975).
  - [4] M. V. Sazhin, *Soviet Ast.* **22**, 36 (1978).
  - [5] S. Detweiler, *ApJ* **234**, 1100 (1979).
  - [6] A. Sesana, A. Vecchio, and C. N. Colacino, *MNRAS* **390**, 192 (2008), 0804.4476.
  - [7] H. Wahlquist, *General Relativity and Gravitation* **19**, 1101 (1987).
  - [8] V. Corbin and N. J. Cornish, *ArXiv e-prints* (2010), 1008.1782.
  - [9] J. A. Ellis, X. Siemens, and J. D. E. Creighton, *ApJ* **756**, 175 (2012), 1204.4218.
  - [10] R. van Haasteren and Y. Levin, *MNRAS* **428**, 1147 (2013), 1202.5932.
  - [11] F. Feroz, M. P. Hobson, E. Cameron, and A. N. Pettitt, *ArXiv e-prints* (2013), 1306.2144.
  - [12] P. Graff, F. Feroz, M. P. Hobson, and A. N. Lasenby, *ArXiv e-prints* (2013), 1309.0790.
  - [13] J. Ellis, *ArXiv e-prints* (2013), 1305.0835.
  - [14] J. P. W. Verbiest, J. M. Weisberg, A. A. Chael, K. J. Lee, and D. R. Lorimer, *ApJ* **755**, 39 (2012), 1206.0428.
  - [15] *ATNF pulsar catalogue*, URL <http://www.atnf.csiro.au/research/pulsar/psrcat>.
  - [16] *PAL: PTA Algorithm Library*, URL <https://github.com/jellis18/PAL>.
  - [17] S. J., in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, edited by T. U. V. Fischer R., Preuss R. (Am. Inst. Phys., New York, 2004), vol. 735, p. 395.
  - [18] F. Feroz and M. P. Hobson, *MNRAS* **384**, 449 (2008), 0704.3704.
  - [19] F. Feroz, M. P. Hobson, and M. Bridges, *MNRAS* **398**, 1601 (2009), 0809.3437.
  - [20] J. Simon, A. Lommen, B. Stappers, L. Finn, F. Jenet, and A. Polin (in preparation).
  - [21] J. A. Ellis, X. Siemens, and R. van Haasteren, *ApJ* **769**, 63 (2013), 1302.1903.
  - [22] P. B. Demorest, R. D. Ferdman, M. E. Gonzalez, D. Nice, S. Ransom, I. H. Stairs, Z. Arzoumanian, A. Brazier, S. Burke-Spolaor, S. J. Chamberlin, et al., *ApJ* **762**, 94 (2013), 1201.6641.
  - [23] A. Whalen, *Detection of signals in noise*, Electrical science series (Academic Press, 1971).
  - [24] P. Jaranowski and A. Królak, *Classical and Quantum Gravity* **27**, 194015 (2010), 1004.0324.
  - [25] R. Prix and B. Krishnan, *Classical and Quantum Gravity* **26**, 204013 (2009), 0907.2569.