COBRA

DOI:
10.1093/mnras/stx2724

Document Version
Final published version

Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **473,** 5026–5042 (2018)
Advance Access publication 2017 October 24

doi:10.1093/mnras/stx2724

COBRA: a Bayesian approach to pulsar searching

L. Lentati,[1⋆] D. J. Champion,[2] M. Kramer,[2,3] E. Barr[2] and P. Torne[2]

[1]*Astrophysics Group, Cavendish Laboratory, JJ Thomson Avenue, Cambridge CB3 0HE, UK*
[2]*Max Planck Institute for Radio Astronomy, Auf dem Hügel 69, D-53121 Bonn, Germany*
[3]*Jodrell Bank Centre for Astrophysics, School of Physics and Astronomy, The University of Manchester, Manchester M13 9PL, UK*

Accepted 2017 October 16. Received 2017 October 16; in original form 2017 March 28

ABSTRACT

We introduce COBRA, a GPU-accelerated Bayesian analysis package for performing pulsar searching, that uses candidates from traditional search techniques to set the prior used for the periodicity of the source, and performs a blind search in all remaining parameters. COBRA incorporates models for both isolated and accelerated systems, as well as both Keplerian and relativistic binaries, and exploits pulse phase information to combine search epochs coherently, over time, frequency or across multiple telescopes. We demonstrate the efficacy of our approach in a series of simulations that challenge typical search techniques, including highly aliased signals, and relativistic binary systems. In the most extreme case, we simulate an 8 h observation containing 24 orbits of a pulsar in a binary with a 30 $M_\odot$ companion. Even in this scenario we show that we can build up from an initial low-significance candidate, to fully recovering the signal. We also apply the method to survey data of three pulsars from the globular cluster 47Tuc: PSRs J0024−7204D, J0023−7203J and J0024−7204R. This final pulsar is in a 1.6 h binary, the shortest of any pulsar in 47Tuc, and additionally shows significant scintillation. By allowing the amplitude of the source to vary as a function of time, however, we show that we are able to obtain optimal combinations of such noisy data. We also demonstrate the ability of COBRA to perform high-precision pulsar timing directly on the single pulse survey data, and obtain a 95 per cent upper limit on the eccentricity of PSR J0024−7204R of $\epsilon_b < 0.0007$.

Key words: methods: data analysis – pulsars: general – pulsars: individual: PSRs J0024−7204D, J0023−7203J and J0024−7204R.

1 INTRODUCTION

Formed from the inevitable gravitational collapse that comes when a giant star runs out of nuclear fuel, neutron stars are amongst the most extreme objects that exist in the Universe. Their large (10^8–10^{15} G) magnetic fields accelerate electrons in the magnetosphere of the neutron star to relativistic velocities, producing beams of radiation that are emitted from above their magnetic poles. When the rotational axis and the magnetic axis of the neutron star are misaligned, then this beam of radiation is swept through space in much the same way as the cone of light from a light-house. If this beam of radiation passes over the Earth then the pulsar becomes visible to us as a regular pulsing beacon in space (see e.g. Lorimer & Kramer 2005 for a detailed overview of the properties of pulsars).

Since the discovery of the first pulsar almost 50 yr ago (Hewish et al. 1968), over 2000 pulsars have been discovered,[1] with recent surveys having been carried out from 100 MHz (Coenen et al. 2014) up to gamma-ray energies (Abdo et al. 2009). More than 10 times this number, however, are expected to be observable in our Galaxy (e.g. Smits et al. 2009).

The study of these extraordinary objects has resulted in a host of groundbreaking discoveries and exciting scientific research. This includes the first evidence for the existence of gravitational waves (Taylor & Weisberg 1989), and the first detection of a planet orbiting another star (Wolszczan & Frail 1992). Pulsars allow us to map the density of the plasma that separates the stars in our Galaxy (e.g. Keith et al. 2013), they provide insights into the physics of star formation and binary evolution (e.g. Tauris & Savonije 1999) and they can even be used to measure the mass of the planets in the Solar system (Champion et al. 2010).

Clearly, the more pulsars that are discovered the greater the potential for scientific progress, with systems like the double pulsar providing high-precision tests of general relativity (GR; Kramer et al. 2006), and 'holy-grail' systems, such as a pulsar in orbit around a black hole, potentially allowing for further tests of gravity, including GR's 'no-hair' theorem (Liu et al. 2014).

Pulsars are known to have steep spectral indices ($S_\nu \propto \nu^{-1.8}$ on average, Maron et al. 2000), and many pulsars have been

⋆ E-mail: lindleylentati@gmail.com
[1] http://www.atnf.csiro.au/people/pulsar/psrcat

observed to have spectra that turn over at lower frequencies, peaking in the 100–200 MHz band (Malofeev, Malov & Shchegoleva 2000) which means they will appear bright in low-frequency surveys. Despite this, the majority of large-scale surveys for new pulsars are performed at approximately 1400 MHz, e.g. the PALFA survey (Cordes et al. 2006) and the HTRU survey (Keith et al. 2010; Barr et al. 2013). This is because the sky temperature also increases as one goes to lower frequencies, and the impact of scattering in the interstellar medium (ISM), which has the effect of broadening the pulse profile, also becomes extremely important at low frequencies. Scattering in particular presents a significant challenge to existing search techniques (e.g. PRESTO, Ransom 2011; SIGPROC, Lorimer 2011), as it spreads out the power in the signal.

While going to much higher frequencies ($\gg$2 GHz) will largely mitigate the effects of the ISM, pulsars tend to be much fainter at such frequencies, and it becomes increasingly challenging to distinguish real candidates from local radio-frequency-interference (RFI). This is because the frequency-dependent delay induced by the ISM actually provides one of the most stringent tests for separating out local from celestial signals.

One location of particular interest in pulsar searching are globular clusters (GCs), which are known to have up to a factor 100 increased density of millisecond pulsars (MSPs) compared to the Galactic plane (e.g. Ransom et al. 2005). In total, 146 radio pulsars have been found in GCs, accounting for more than 5 per cent of the total known pulsar population and a large fraction of the fastest rotating MSPs. *Fermi* gamma-ray observations, however, suggest that significantly larger numbers of yet undetected MSPs could still reside in these clusters (Abdo et al. 2010). Simulations have also shown that Galactic GCs could house between 600 and 3700 pulsars (Turk & Lorimer 2013). GC searches also have the advantage of being the target of repeated pointings made over many years, although this is an advantage that is poorly leveraged by the community, with only simple procedures having been used to incoherently combine different epochs (Sulman, Ransom & Stinebring 2005; Aulbert 2007; Pan et al. 2016, see also Lorimer & Kramer 2005 and references therein), or by performing semicoherent searches were information about candidates from coherent searches of short segments of data are combined afterwards incoherently (e.g. Smith 2016).

In this paper, we will introduce COBRA, which provides a fundamentally different approach to performing a search for new pulsars. An outline of the analysis method used is shown in Fig 1. First, COBRA does not perform any analysis using the power spectrum. Instead, we construct a pulse train using a physically motivated timing model and a simple Gaussian profile, which is then compared directly to the search data. COBRA therefore makes full use of phase information in the data set, which makes it possible to coherently combine multiple epochs that may be separated by weeks or months, in order to maximize the sensitivity of the search to the pulsar without the need for padding extremely large Fourier transforms. By constructing a model for the pulse train, we can directly incorporate the effect of (i) pulse-broadening due to scattering in the ISM, (ii) aliasing of the signal that results from a pulse-width that is narrower than the sampling time-scale of the survey data and (iii) the effect of a binary companion on the arrival time of the pulses.

COBRA, however, is not a blind-search technique – the prospect of exploring the full pulsar parameter space (i.e. from approximately 1 ms to 10 s) through Bayesian methods is still impractical computationally. We therefore make use of existing tools, such as SIGPROC[2] and PRESTO,[3] to first produce a list of candidates for each epoch in the survey, and use the period and DM of these candidates to restrict our prior for these parameters in the Bayesian search. Using COBRA, we can set a threshold for the signal to noise (S/N) of candidates from each epoch separately, such that if that source were present in all the data at a similar significance, then a coherent analysis of all the survey data would result in a detection. Although this requires that existing tools return some small hint of a signal, we will show that from that point onwards, COBRA then enables superior modelling of effects such as scintillation, aliasing, scattering and relativistic binary (RB) effects, not only when there have been repeated observations, but also using only single pointings.

In Section 2, we briefly introduce Bayesian analysis methods, and relate them to the Frequentist methods typically used in pulsar searching. In Section 3, we describe our approach to Bayesian pulsar searching, including the models used for scattering, and to incorporate the effects of both Keplerian and RB orbits. We then apply this method to simulations in Section 4, briefly consider optimizations of the Graphical Processing Unit (GPU) likelihood in Section 5 and then analyse several known pulsars in the GC 47Tuc in Section 6. Finally, in Section 7, we describe how COBRA can use phase information to better discriminate against sources of RFI, before we offer some concluding remarks in Section 8.

2 BAYESIAN ANALYSIS AND MODEL SELECTION

For a complete discussion of Bayesian methods, see for example, Sivia & Skilling (2006). Here, we will only introduce some basic terminology, and draw some simple comparisons with Frequentist methods, as these have thus far dominated in pulsar searching.

When performing a Frequentist analysis, the quantity of interest is typically the likelihood, $\mathcal{L}(\theta) = \Pr(\mathcal{D} \,|\, \theta, \mathcal{H})$, which is the probability of our data set, $\mathcal{D}$, given our model parameters, θ, in a model or hypothesis $\mathcal{H}$. Here then, the model parameters are assumed to be a fixed quantity, and the observed data is assumed to be a random variable. In a simple case, where our data set is a single point, d, subject to known Gaussian noise σ, the likelihood that our data is described by some model parameter m will be given by

$$\Pr(d \,|\, m) = \exp\left(\frac{(d-m)^2}{\sigma^2}\right). \tag{1}$$

In Bayesian analysis, we are interested in the inverse of this quantity, referred to as the posterior distribution $\mathcal{P}(\theta)$. The posterior is the probability of our parameters, given our observed data and our model, i.e. $\mathcal{P}(\theta) = \Pr(\theta \,|\, \mathcal{D}, \mathcal{H})$. Here then, we assume the data is a fixed quantity, and our model parameters are the random variables. The posterior can be obtained using Bayes equation:

$$\Pr(\theta \,|\, \mathcal{D}, \mathcal{H}) = \frac{\Pr(\mathcal{D} \,|\, \theta, \mathcal{H})\Pr(\theta \,|\, \mathcal{H})}{\Pr(\mathcal{D} \,|\, \mathcal{H})}. \tag{2}$$

The term $\Pr(\theta \,|\, \mathcal{H})$ is referred to as the prior probability distribution of our parameters. These are often uniform in the parameter, or the log of the parameter over some range, but can also be more constraining if there is more prior information available. The final term in Bayes equation, $\Pr(\mathcal{D} \,|\, \mathcal{H})$ is referred to as the Evidence, and is critical in performing selection between different hypotheses.

[2] http://sigproc.sourceforge.net

[3] http://www.cv.nrao.edu/~sransom/presto

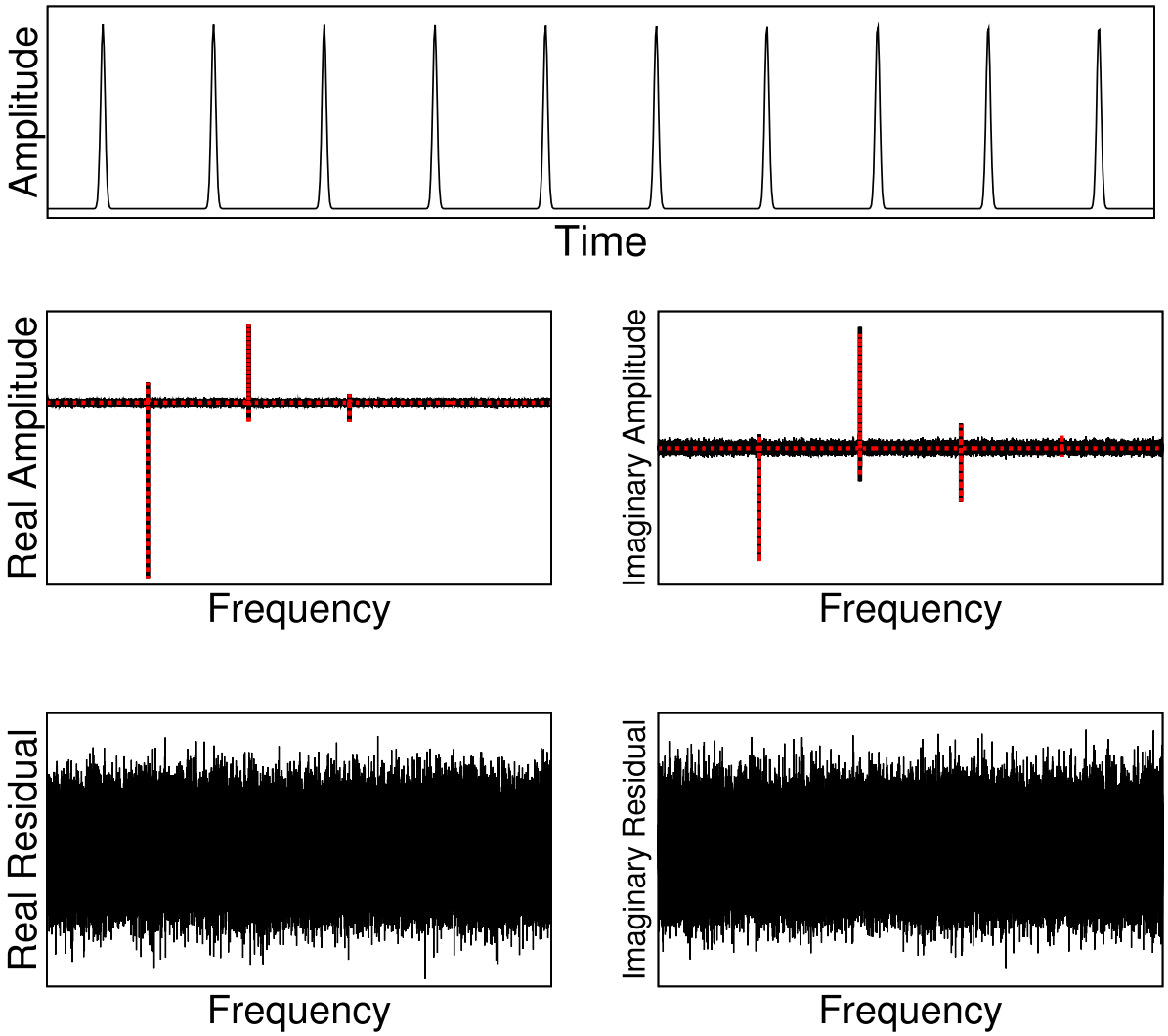

Figure 1. From top to bottom, an overview of the three main steps in the analysis process that takes place using COBRA. (1) We first build a pulse train in time domain assuming a Gaussian profile, and a physical timing model. (2) This pulse train is then Fourier transformed (red lines) and compared to the Fourier transformed data (black lines). (3) The residuals are obtained by subtracting the model from the data, and the likelihood of the model is evaluated using equation (54).

The evidence is given by the integral of the likelihood over the full parameter space, weighted by the prior:

$$\Pr(\mathcal{D} \mid \mathcal{H}) = \int \Pr(\mathcal{D} \mid \theta, \mathcal{H}) \Pr(\theta \mid \mathcal{H}) \mathrm{d}^n\theta, \tag{3}$$

where n is the number of parameters in the model. The Evidence automatically implements Occam's razor, in that a simpler model that spans a smaller volume of parameter space will have a larger Evidence than a more complex model, unless that more complex model is able to describe the data significantly better.

Thus, while for Frequentist methods we would be interested in the number of templates that must be searched over in order to construct a measure of significance, in Bayesian analysis we compute the Evidences for two models and use them to obtain their relative probabilities. While in principle we could discretize our parameter space as in the case of a template bank, in all the following work we vary the parameters smoothly within the prior range quoted, unless explicitly stated. For example, if we have the two scenarios:

$\mathcal{H}_1$: There is a pulsar in the data set

$\mathcal{H}_2$: There is no pulsar in the data set

we can compute the Evidence for each model, and calculate the 'odds ratio' of $\mathcal{H}_1$ compare to $\mathcal{H}_2$ directly, via:

$$R = \frac{\Pr(\mathcal{D} \mid \mathcal{H}_1)}{\Pr(\mathcal{D} \mid \mathcal{H}_2)}. \tag{4}$$

An odds ratio of <1 would imply no support for the presence of a pulsar in the data set, while an odds ratio of $\gtrsim 20$ would begin to suggest significant support (see e.g. Kass & Raftery 1995 for more information on the interpretation of odds ratios).

In the following, rather than discuss changes in the Evidence itself, we will be computing the log Evidence, and thus will quote differences in this quantity. Therefore, an increase in the log Evidence for $\mathcal{H}_1$ of three, is considered significant.

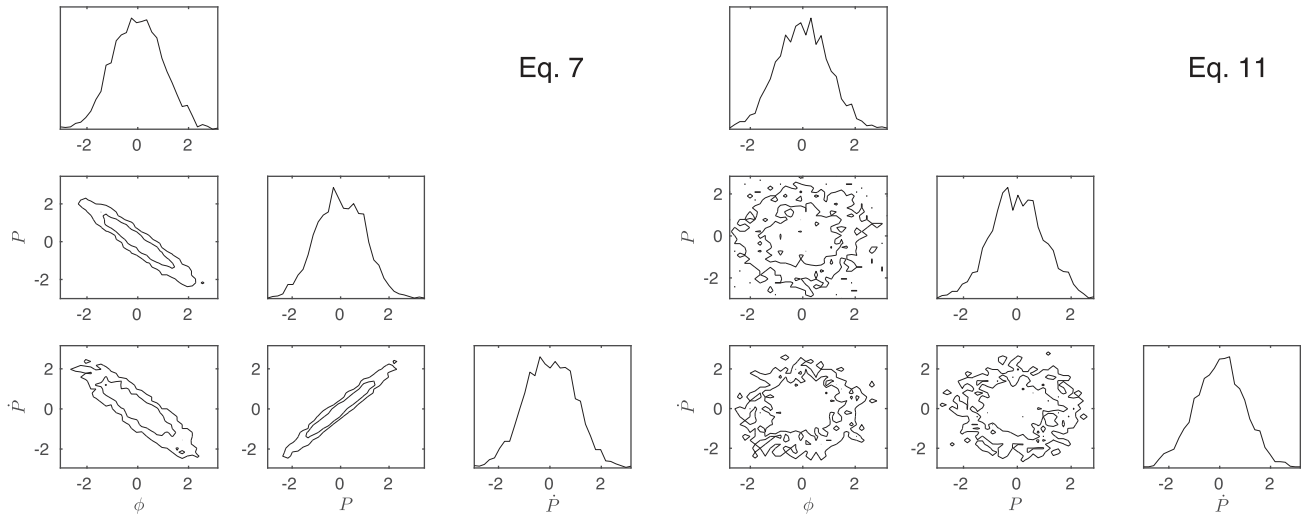

Figure 2. One- and two-dimensional posterior distributions for phase, period and period derivative when using equation (7) (left-hand panel), and when using equation (11) (right-hand panel) as defined in Section 3. Posteriors are shown after subtracting the mean, and scaling by the standard deviation of the distribution. By decorrelating the parameter space, we decrease the number of samples required (and thus the run time) by a factor of 3.

3 PULSAR SEARCHING LIKELIHOOD

We first consider our data, $\boldsymbol{d}$, to be a single de-dispersed time series consisting of N_d data points, d_i, observed at a topocentric time t_i, where henceforth the subscript i will refer to a specific index in the associated vector. We then calculate the Fourier transform of the data, $\tilde{\boldsymbol{d}}$, which will contain $N_{\rm d}/2 + 1$ real and imaginary values, and discard both the dc-offset, and nyquist terms, leaving us with $N_{\rm d}/2 - 1$ complex values. As henceforth we will only refer to the Fourier representation of the data, we will simply denote it $\boldsymbol{d}$.

While our time series is defined in the topocentre, our model for the pulse train will be defined at the Solar system Barycenter (SSB), and so we must obtain the transformation that takes our sampled times from the topocentre to the SSB. We obtain this transformation using the PYTHON interface to the TEMPO2 timing package, LIBSTEMPO.[4] For each time series we construct a fake data set which includes: (i) a simple timing ephemeris that includes the true sky position of the pointing, and an arbitrary period and dispersion measure and (ii) a series of site arrival times spaced at intervals of 10 s over the observation period. LIBSTEMPO then directly returns the required conversion. Note that in doing this correction to the SSB we are at no point modifying the data, as opposed to PRESTO and SIGPROC, which alters the time series data itself in order to have it match the barycentered time. In COBRA, this transformation is only applied to the time stamps $\boldsymbol{t}$, and is used to adjust the model pulse train.

3.1 Isolated and weakly-accelerated pulsars

In the simplest model, we will include only a model period P, a phase offset ϕ and the width, $\sigma_{\rm p}$, of a Gaussian pulse shape used to construct the pulse train. We first adjust the barycentered times, $\boldsymbol{b}$, using the phase offset to compute the *signal times*, $\boldsymbol{s}$:

$$s(\phi)_i = b_i - \phi. \tag{5}$$

We then transform these times to the range $-\frac{P}{2}$ to $\frac{P}{2}$ to obtain the phase-wrapped time coordinates $\bar{\boldsymbol{s}}$:

$$\bar{s}_i = \frac{s_i}{P} - \mathrm{trunc}\left(\frac{s_i}{P}\right) - \mathrm{trunc}\left(\frac{s_i}{P} - \mathrm{trunc}\left(\frac{s_i}{P} + 1\right)\right) - 0.5. \tag{6}$$

In order to include a period derivative, $\dot{P}$, in the model, we then simply alter equation (5):

$$s(\phi, \dot{P})_i = b_i - \phi - \dot{P} b_i^2. \tag{7}$$

Written as in equation (7) the phase, period and period derivative parameters have the potential to be extremely correlated, which can significantly decrease the sampling efficiency, increasing the time it takes to evaluate the model for a particular candidate.

We can, however, project out a large fraction of this covariance. We first consider the phase and period parameters, which we can decorrelate by adding the term:

$$\delta\phi_{\rm p} = \frac{1}{P}\left(\frac{T}{2} + b_0\right) \tag{8}$$

to the phase offset, where b_0 is the barycentered time of the first bin in the time series. For the period derivative we need to consider the contribution to both the phase offset, $\delta\phi$, given by

$$\delta\phi_{\dot{P}} = \int_{b_0}^{b_N} \dot{P} x_i^2 \mathrm{d}x = \frac{\dot{P}}{3}\left(b_N^3 - b_0^3\right), \tag{9}$$

where b_N is the barycentered time of the last bin in the time series, and to the period, $\delta P_{\dot{P}}$:

$$\delta P_{\dot{P}} = \frac{\dot{P}(b_N^2 - b_0^2)}{T}. \tag{10}$$

We can then simply rewrite equation (7) as

$$s(\phi, P, \dot{P})_i = b_i - \phi - \dot{P} b_i^2 + \delta\phi_P + \delta\phi_{\dot{P}} + \delta P_{\dot{P}}(b_o - b_i), \tag{11}$$

where b_o is the weighted mean of the barycentered arrival times. In Fig. 2, we compare the one- and two-dimensional posterior distributions for phase, period and period derivative when using equation (7) (left-hand panel), and when using equation (11) (right-hand panel). By decorrelating the parameter space, we decrease the number of

[4] https://github.com/vallis/libstempo

samples required (and thus the run time) by a factor of approximately three.

3.2 Binary models

In the following section, we describe our implementation of four different binary models:

(i) A three-parameter circular binary (CB), which requires a binary amplitude, $A_{\rm b}$, binary phase, $\phi_{\rm b}$ and binary period, $P_{\rm b}$.
(ii) A five-parameter eccentric binary (EB), for which we include the eccentricity, $\epsilon_{\rm b}$, and longitude of periastron, $\omega_{\rm b}$.
(iii) A seven-parameter RB, which assumes GR, and in addition to the EB model requires the pulsar mass $M_{\rm p}$, and the companion mass $M_{\rm c}$.
(iv) A six- to ten-parameter post-Keplerian (PK) binary (PKB), in which we directly include the PK terms (see Section 3.2.3).

3.2.1 Circular binaries

For the CB model, the modification to the barycentric times, $\boldsymbol{b}$, is given by Damour & Deruelle (1986), henceforth DD:

$$B_{\rm c}(A_{\rm b}, \phi_{\rm b}, P_{\rm b})_i = C_1\,(1 - C_2 + C_3)\,, \tag{12}$$

where

$$C_1(A_{\rm b}, \phi_{\rm b}, P_{\rm b})_i = A_{\rm b} \sin\left(\frac{2\pi b_i}{P_{\rm b}} + \phi_{\rm b}\right), \tag{13}$$

$$C_2(A_{\rm b}, \phi_{\rm b}, P_{\rm b})_i = \frac{2\pi}{P_{\rm b}} A_{\rm b} \cos\left(\frac{2\pi b_i}{P_{\rm b}} + \phi_{\rm b}\right), \tag{14}$$

$$C_3(A_{\rm b}, \phi_{\rm b}, P_{\rm b})_i = \frac{1}{2}\left[\frac{2\pi}{P_{\rm b}} A_{\rm b} \sin\left(\frac{2\pi b_i}{P_{\rm b}} + \phi_{\rm b}\right)\right]^2. \tag{15}$$

When only sampling a small fraction of a binary orbit, as with the period derivative in Section 3.1, the binary parameters become extremely covariant with the overall phase offset and period of the pulsar. Approximating equation (12) as only the C_1 term, we analytically compute the contribution of the binary signal to the phase offset, $\delta\phi_{\rm c}$, and the period, $\delta P_{\rm c}$:

$$\delta\phi_{\rm c} = \frac{P_{\rm b}}{2T\pi}\left[\cos\left(\frac{2\pi b_0}{P_{\rm b}} + \phi_{\rm b}\right) - \cos\left(\frac{2\pi b_N}{P_{\rm b}} + \phi_{\rm b}\right)\right], \tag{16}$$

and

$$\delta P_{\rm c} = \frac{1}{T}\left[\sin\left(\frac{2\pi b_N}{P_{\rm b}} + \phi_{\rm b}\right) - \sin\left(\frac{2\pi b_0}{P_{\rm b}} + \phi_{\rm b}\right)\right]. \tag{17}$$

Further, the amplitude of the binary signal is extremely covariant with the other binary parameters, all of which can significantly decrease the sampling efficiency. To mitigate this, we subtract $\delta\phi_{\rm c}$ and $\delta P_{\rm c}$ from the binary signal, and then normalize it by its standard deviation. While this correlates the binary amplitude with the phase and period parameters, we find this to be more efficient than sampling from the binary amplitude directly.

In order to avoid evaluating trigonometric functions on the GPU, we interpolate $B_{\rm c}$ as a function of binary phase. For a CB this is straight forward, as the interpolation is independent of all three binary parameters. We therefore pre-compute 100 000 uniformly sampled points for $\sin(\theta)$ and $\cos(\theta)$, with θ between zero and 2π, and linearly interpolate between these values in the likelihood function. In Fig. 3 (top), we show the difference between the interpolated function and the true binary model as a function of binary phase.

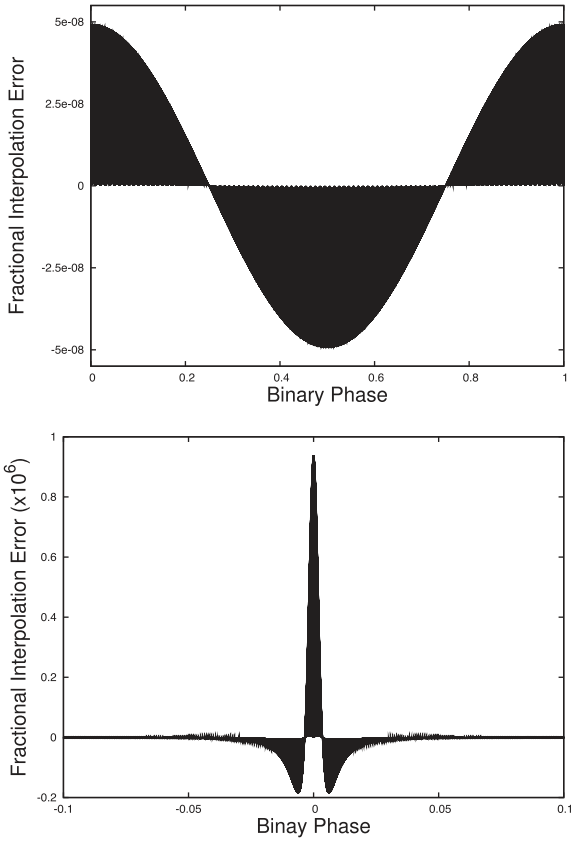

Figure 3. Fractional errors introduced as a result of interpolating the circular (top) and eccentric ($\epsilon = 0.9$, bottom) binary orbit in phase.

For a binary signal with a maximum amplitude of 1 s, the largest deviation is less than 50 ns, which is sufficient for the purposes of a search.

3.2.2 Eccentric binaries

The EB system requires that we solve Kepler's Equation for the eccentric anomaly of the orbit. Given the mean anomaly, $\mathcal{M}$, which for barycentric time b_i is defined as

$$\mathcal{M}_i = \frac{2\pi b_i}{P_{\rm b}} + \phi_{\rm b}, \tag{18}$$

we can arrive at the eccentric anomaly, $\mathcal{E}$ by iterating with the Newton–Raphson technique. This requires an initial guess of $\mathcal{E}$, which we take to be

$$\mathcal{E}_0 = \mathcal{M}_i + \frac{\epsilon_{\rm b} \sin(\mathcal{M}_i)}{\sqrt{1 - 2\epsilon_{\rm b}\cos(\mathcal{M}_i) + \epsilon_{\rm b}^2}} \tag{19}$$

and uses the gradient:

$$\mathrm{d}\mathcal{E}_i = \frac{\mathcal{E}_i - \epsilon_{\rm b}\sin(\mathcal{E}_i) - \mathcal{M}_i}{1 - \epsilon_{\rm b}\cos(\mathcal{E}_i)} \tag{20}$$

to converge on $\mathcal{E}$ within a few iterations.

Following DD we then define the following quantities:

$$\eta = \frac{2\pi}{P_{\rm b}(1 - \epsilon_{\rm b}\cos(\mathcal{E}))}, \tag{21}$$

$$\alpha = A_{\rm b}\sin(\omega_{\rm b}), \tag{22}$$

$$\beta = A_{\rm b}\sqrt{1-\epsilon_{\rm b}^2}\cos(\omega_{\rm b}), \tag{23}$$

$$E_1 = \alpha(\cos(\mathcal{E}) - \epsilon_{\rm b}) + \beta\sin(\mathcal{E}), \tag{24}$$

$$E_2 = -\alpha\sin(\mathcal{E}) + \beta\cos(\mathcal{E}), \tag{25}$$

$$E_3 = -\alpha\cos(\mathcal{E}) - \beta\sin(\mathcal{E}), \tag{26}$$

and write the modification to the barycentric times due to the EB as

$$B_{\rm e} = E_1\left(1 - \eta E_2 + \eta^2\left(E_2^2 + \frac{1}{2}E_1E_3 - \frac{1}{2}\frac{\epsilon_{\rm b}\sin(\mathcal{E})E_1E_2}{1-\epsilon_{\rm b}\cos(\mathcal{E})}\right)\right). \tag{27}$$

As with the circular orbit we compute the contribution of the orbit to the period and phase parameters in the model. In this case, however, we must do this numerically as opposed to analytically. We also pre-compute $\cos(\mathcal{E})$ and $\sin(\mathcal{E})$ for a series of different eccentricities. When first including the eccentric model in an analysis, we pre-compute our eccentricity grid in steps of 0.01 up to a maximum of 0.99. This grid can then be refined in subsequent analyses as the constraints improve. In Fig. 3 (bottom), we show the deviation from the true orbit for an eccentricity of 0.9 where, as for the circular orbit, we interpolate in binary phase in 100 000 steps. The maximum error is less than a microsecond, which given the typical sampling time of search data of tens of μs, is still over an order of magnitude less than one sample bin.

3.2.3 Relativistic binaries

Finally, we consider relativistic binaries, for which we include several additional PK corrections to the change in arrival time, which describe:

(i) $\dot{P}_{\rm b}$ – first time derivative of the binary period,
(ii) $\dot{\omega}_{\rm b}$ – first time derivative of the longitude of periastron,
(iii) γ – gravitational redshift and time dilation as the pulsar moves around its common centre of mass,
(iv) r, s – Shapiro delay 'range' and 'shape', caused by the gravitational field of the companion,
(v) δ_θ, $\delta_{\rm r}$ – relativistic deformation of the orbit from an ellipse.

Under GR (and indeed more generally, see Damour & Taylor 1991), these contributions can all be derived using just the mass of the pulsar, $M_{\rm p}$, and the mass of the companion, $M_{\rm c}$, as in DD:

$$\dot{P}_{\rm b} = -\frac{192\pi}{5}T_\odot^{5/3}\left(\frac{P_{\rm b}}{2\pi}\right)^{-5/3} f(\epsilon_{\rm b})\frac{M_{\rm p}M_{\rm c}}{(M_{\rm p}+M_{\rm c})^{1/3}}, \tag{28}$$

$$\dot{\omega}_{\rm b} = 3T_\odot^{2/3}\left(\frac{P_{\rm b}}{2\pi}\right)^{-2/3}\frac{1}{1-\epsilon_{\rm b}^2}(M_{\rm p}+M_{\rm c})^{2/3}, \tag{29}$$

$$\gamma = T_\odot^{2/3}\left(\frac{P_{\rm b}}{2\pi}\right)^{1/3}\epsilon_{\rm b}\frac{M_{\rm c}(M_{\rm p}+2M_{\rm c})}{(M_{\rm p}+M_{\rm c})^{4/3}}, \tag{30}$$

$$r = T_\odot M_{\rm c}, \tag{31}$$

$$s = T_\odot^{-1/3}\left(\frac{P_{\rm b}}{2\pi}\right)^{-2/3}A_{\rm b}\frac{(M_{\rm p}+M_{\rm c})^{2/3}}{M_{\rm c}}, \tag{32}$$

$$\delta_\theta = T_\odot^{2/3}\left(\frac{P_{\rm b}}{2\pi}\right)^{-2/3}\frac{\frac{7}{2}M_{\rm p}^2 + 6M_{\rm p}M_{\rm c} + 2M_{\rm c}^2}{(M_{\rm p}+M_{\rm c})^{4/3}}, \tag{33}$$

$$\delta_{\rm r} = T_\odot^{2/3}\left(\frac{P_{\rm b}}{2\pi}\right)^{-2/3}\frac{3M_{\rm p}^2 + 6M_{\rm p}M_{\rm c} + 2M_{\rm c}^2}{(M_{\rm p}+M_{\rm c})^{4/3}}, \tag{34}$$

where $T_\odot = 4.925490947$ μs, and

$$f(\epsilon_{\rm b}) = \frac{1 + (73/24)\epsilon_{\rm b}^2 + (37/96)\epsilon_{\rm b}^4}{(1-\epsilon_{\rm b}^2)^{7/2}}. \tag{35}$$

While the impact of the relativistic deformation parameters ($\delta_{\rm r}$, δ_θ) is extremely small, with measurements only now being made for a small number of systems (Weisberg & Huang 2016; Kramer et al. in preparation), under the assumption of GR they can be included at practically zero computational cost, and so we use them in our RB model regardless. In our PKB binary model, any of the PK terms listed above can be included as free parameters, however the Shapiro delay, s must be included in combination with the companion mass $M_{\rm c}$.

We note that in addition to those terms listed above, there are further 'aberration delays' due to the difference between the proper time of the pulse emission, and the emission time if the pulsar were not a rotating body. However, as discussed in DD, these delays can be absorbed by other orbital parameters, and so unless the binary undergoes geodetic precession, they cannot be isolated in the timing model. We therefore do not currently implement them in our binary model.

As for the EB, we must solve Kepler's equation for the eccentric anomaly. We do this as in Section 3.2.2, however, in this case, the definition of the mean anomaly, $\mathcal{M}$, includes a contribution from the period derivative, $\dot{P}_{\rm b}$:

$$\mathcal{M}_i = \frac{2\pi b_i}{P_{\rm b}} + \phi_{\rm b} - \frac{1}{2}\frac{\dot{P}_{\rm b}}{2\pi}\left(\frac{2\pi b_i}{P_{\rm b}} + \phi_{\rm b}\right)^2. \tag{36}$$

Additionally, the change in $\omega_{\rm b}$ as a function of time means that we must now also compute the true anomaly, $\mathcal{T}$, via:

$$\mathcal{T}_i = 2\arctan\left[\sqrt{\frac{1+\epsilon_{\rm b}}{1-\epsilon_{\rm b}}}\tan\left(\frac{\mathcal{E}_i}{2}\right)\right], \tag{37}$$

so that the value of $\omega_{\rm b}$ at a particular barycentric time, i, is given by

$$\omega = \omega_{\rm b} + \dot{\omega}_{\rm b}\mathcal{T}_i. \tag{38}$$

Using the above PK terms, and the expressions for ω and $\mathcal{E}$, we can then define the following quantities:

$$\epsilon_{\rm r} = \epsilon_{\rm b}(1+\delta_{\rm r}), \tag{39}$$

$$\epsilon_\theta = \epsilon_{\rm b}(1+\delta_\theta), \tag{40}$$

$$\eta = \frac{2\pi}{P_{\rm b}(1-\epsilon_{\rm b}\cos(\mathcal{E}))}, \tag{41}$$

$$\alpha = A_{\rm b}\sin(\omega), \tag{42}$$

$$\beta = A_{\rm b}\sqrt{1-\epsilon_\theta^2}\cos(\omega) + \gamma, \tag{43}$$

$$G_1 = \alpha(\cos(\mathcal{E}) - \epsilon_r) + \beta \sin(\mathcal{E}), \tag{44}$$

$$G_2 = -\alpha \sin(\mathcal{E}) + \beta \cos(\mathcal{E}), \tag{45}$$

$$G_3 = -\alpha \cos(\mathcal{E}) - \beta \sin(\mathcal{E}), \tag{46}$$

$$G_4 = -2r \ln \Big[1 - \epsilon_b \cos(\mathcal{E}) - s\,(\sin(\omega)(\cos(\mathcal{E}) - \epsilon_b) + \sqrt{1-\epsilon^2}\cos(\omega)\sin(\mathcal{E})\Big)\Big], \tag{47}$$

so that the final modification to our barycentric arrival times due to the RB motion is given by

$$B_{gr} = G_1\left(1 - \eta G_2 + \eta^2 \times \left(G_2^2 + \frac{1}{2}G_1G_3 - \frac{1}{2}\frac{\epsilon_b \sin(\mathcal{E}) G_1 G_2}{1-\epsilon_b\cos(\mathcal{E})}\right)\right) + G_4. \tag{48}$$

For the relativistic model, in addition to pre-computing $\cos(\mathcal{E})$ and $\sin(\mathcal{E})$, we also store the true anomaly for each of the eccentricities in the interpolation grid.

3.3 Forming the pulse train

Given a signal model for the arrival times we then assume a Gaussian model for the pulse profile with a standard deviation denoted by σ_p. We can therefore construct our model pulse train by

$$\boldsymbol{m}(\bar{s}, \sigma_p) = \exp\left[-\frac{1}{2}\bar{s}^2/\sigma_p^2\right], \tag{49}$$

which we FFT to get our Fourier domain model $\tilde{\boldsymbol{m}}$. As with the data vector, as we will only deal with the Fourier representation of the signal we drop the tilde and denote the Fourier domain model simply as $\boldsymbol{m}$, and as before, we remove the dc-offset and Nyquist terms. Note that here we explicitly assume that the pulse profile is a Gaussian. As such any non-Gaussianity in the pulse shape will result in a mismatch between the data and the model, and thus will reduce the overall S/N of the detection. We explore this loss in signal in Section 6 in the context of PSR J0024−7204R.

3.4 Scattering

The effect of scattering due to the ISM is to broaden the observed pulse profile by convolving the intrinsic profile with a pulse broadening function (PBF). In the Fourier domain, this convolution is therefore simply a multiplication of our model signal, with the Fourier transform of the model PBF.

We assume a simple thin screen model for the PBF (Williamson 1972), using a single parameter for the time-scale, τ, which in the time domain will be given by

$$\mathrm{PBF}(t, \tau, \nu) = H(t)\exp\left(-\frac{t}{\bar{\nu}^{-4}\tau}\right). \tag{50}$$

where $\bar{\nu} = \nu/\nu_0$ is the observing frequency, ν, normalized by an arbitrary reference frequency, ν_0, and $H(t)$ is the Heavyside step function. The analytic Fourier transform of equation (50) is then given by

$$\mathrm{PBF}(f, \tau, \nu) = \frac{1}{(2\pi f \bar{\nu}^{-4}\tau)^2 + 1} + \frac{-2\pi f \bar{\nu}^{-4}\tau}{(2\pi f \bar{\nu}^{-4}\tau)^2 + 1} i, \tag{51}$$

where i indicates a complex number. We can multiply the Fourier representation of our signal, $\boldsymbol{m}$, by equation (51) to get the scattered signal model.

3.5 Marginalizing over the amplitude

We can now write our likelihood as a function of the overall amplitude, A, of our model vector $\boldsymbol{m}$, and the parameters that went into $\boldsymbol{m}$, which for convenience we will collectively refer to as Θ as

$$\Pr(A, \Theta|\boldsymbol{d}) = \exp\left(-\frac{1}{2\sigma_n^2}(\boldsymbol{d} - A\boldsymbol{m})^T(\boldsymbol{d} - A\boldsymbol{m})\right). \tag{52}$$

Here, σ_n is an estimate of the white noise in the data set, for which we obtain an estimate from the high-frequency end of the Fourier data. While this model therefore currently explicitly assumes that the noise in the Fourier data is well described by an uncorrelated, uniform Gaussian process, RFI will mean that this assumption is not true. Ideally one would like to incorporate a model for the RFI directly into the analysis, and achieving this in a computationally efficient manner will be the subject of future work.

Rather than sampling numerically over the overall amplitude parameter, we marginalize over it analytically. We define $\Sigma = \boldsymbol{m}^T\boldsymbol{m}$ and $\bar{d} = \boldsymbol{d}^T\boldsymbol{m}$, and $n = \sqrt{\Sigma/2/N_f}$ so that our marginalized log-likelihood is then given by

$$L(\Theta|\boldsymbol{d}) = \frac{1}{2}\bar{d}\,\Sigma^{-1}\bar{d} - \frac{1}{2}\log\Sigma - \log n. \tag{53}$$

We can now trivially extend equation (53) to incorporate additional epochs or observing frequencies by simply writing it as

$$L(\Theta|\boldsymbol{d}) = \sum_i^{N_d} \frac{1}{2}\bar{d}_i\,\Sigma_i^{-1}\bar{d}_i - \frac{1}{2}\log\Sigma_i - \log n_i, \tag{54}$$

where the sum is over the N_d different time series included in the full analysis.

4 APPLICATION TO SIMULATIONS

In the following subsections, we apply COBRA to a series of simulations designed to challenge existing search techniques. In particular we consider:

(i) Noisy signals (Section 4.1), where the pulsar suffers significant scintillation.
(ii) Highly aliased signals (Section 4.2), where the pulse width is a small fraction of the sampling time of the survey data.
(iii) Binary systems (Section 4.3), for circular, eccentric and relativistic orbits.

In all cases, we produce the simulations using the SIGPROC functions FASTFAKE and INJECTPULSAR. These use TEMPO2 to evaluate the timing model for a particular ephemeris, and produce a time series given that ephemeris and a model for the profile. We specify in each subsection the specific details of the simulations, such as the integrated S/N of the data set which is equivalent to the S/N of the folded profile, the sampling time and observation length.

4.1 Combining noisy data

When previous attempts have been made to combine multiple epochs, the addition is performed incoherently or semicoherently, and the stacking is performed assuming that the amplitude of the signal is the same in every epoch. However, effects such as scintillation, inclement weather or RFI can cause the amplitude of the signal

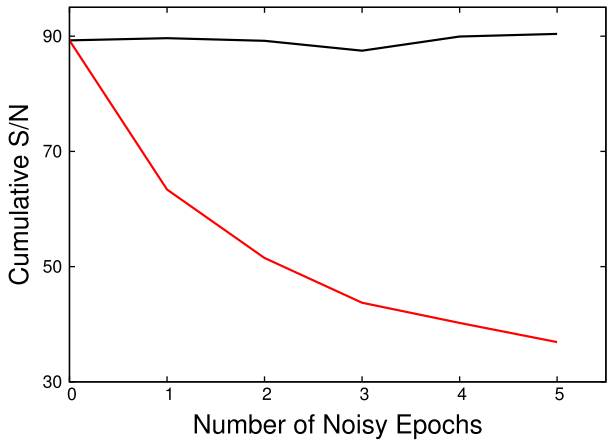

Figure 4. Cumulative S/N when one search epoch that contains a significant detection is combined with an increasing number of noise-only data sets. We compare models that (i) assumes the same amplitude of the signal in each epoch (red line) or (ii) includes a different amplitude in each epoch (black line).

to vary by orders of magnitude. If epochs are added in which the source is experiencing a minimum in the observed brightness, then simply adding this to a high S/N observation under the assumption that the source is equally bright in both, will decrease the overall S/N of the combined detection by $\sim\sqrt{2}$, equivalent to simply adding noise to the original data set.

In equation (54) one can see that, in our framework, different epochs are each assigned an overall amplitude parameter for the signal. As such, epochs that are 'just noise' will be effectively down-weighted, and will not contribute to the overall S/N of the signal. In Fig. 4, we show the cumulative S/N for a significant detection, after adding a series of additional epochs into our analysis that only contain noise (black line). The significance is seen to suffer a slight random walk around that obtained when including no additional epochs. The red line shows the impact of combining noisy epochs when the model assumes the same amplitude in each epoch. In this case, the significance drops at a rate of $\sqrt{N}$, with N the total number of epochs, as expected in this case. By allowing the amplitude to vary in each epoch, we therefore correctly account for our uncertainty in the scintillation properties of the source, and obtain a significantly more optimal result.

4.2 Aliasing

We simulate a simple isolated pulsar with a period of 3.99 ms, using the pulse profile from the MSP J1909−3744, which has a full width at half-maximum (FWHM) of approximately 120 μs. We initially simulate this pulsar with a sampling time, τ_b, of 80 μs, and then down-sample by factors of 2 up to a factor of 32, which gives a sampling time of 2560 μs. The FWHM of the pulse profile is thus sampled from between 1.5 and 0.05 bins. In Fig. 5 (top-left panel), we show the folded pulse profile using the injected period for downsampling factors of 1, 2, 6 and 32. The observed width, σ_o, is given by approximately:

$$\sigma_o = \sqrt{\left(\frac{\tau_b}{2}\right)^2 + \sigma_p^2}\,, \tag{55}$$

with σ_p the intrinsic standard deviation of the pulse profile. We would therefore expect that as we downsample the data, the maximum S/N that we could extract would decrease as $\sqrt{\sigma_p/\sigma_o}$. In Fig. 5 (top-right panel), we show this decrease in the S/N (black line) compared to the S/N extracted by our analysis with COBRA (red line). Even in the most extreme case, we find that out our analysis tracks the expected S/N. In the bottom panel of Fig. 5, we show the real (left-hand panel) and imaginary (right-hand panel) parts of the data and the model for the 80 μs sampled data. In this case, the power is spread across a large number of harmonics, which is problematic for search techniques that rely on 'harmonic summing', where the powers in the N, lowest frequency, harmonics are added together under the assumption of equal weighting. In COBRA, our model for the pulsar is defined across the whole of the sampled Fourier domain, and so we automatically recover all of the signal present in the data. Further, as we include the pulse width as a free parameter in our analysis, we are able to recover any signal that gets aliased to lower frequencies when the data is FFT'ed, as that same FFT is applied to the signal model.

We note that features in the pulse profile that are narrower than the raw sampling rate will typically not manifest as aliasing, as the signal is integrated over the specified time resolution, rather than actually being sampled. However, extremely narrow features of the order ∼10 ns have been observed in pulse profiles (e.g. Soglasnov et al. 2004), and the power in those features will be spread across an extremely large number of harmonics in the Fourier domain.

4.3 Binary systems

In the next three subsections, we simulate data sets that include circular, eccentric and RB systems.

4.3.1 Circular binary

We first simulate 5 min, high S/N, data sets for four CB systems. We use orbital periods of (i) 1 d, (ii) 0.1 d, (iii) 0.05 d and (iv) 0.01 d for the four simulations, respectively. This means we are progressively sampling larger portions of the binary orbit, initially only covering 0.35 per cent of the orbit, and increasing up to 35 per cent. In each case, we use the Bayesian Evidence to perform a comparison between a simple acceleration model, and the full CB model.

As introduced in Section 2, the Evidence is the average of the likelihood over the prior, and so is larger for a model if more of its parameter space is likely and smaller for a model where large areas of its parameter space have low likelihood values, even if the likelihood function is very highly peaked. The Evidence thus automatically implements Occam's razor: a simpler theory with a compact parameter space will have a larger evidence than a more complicated one, unless the latter is significantly better at explaining the data.

We find that for the longest period, the simpler acceleration model is sufficient to describe the data, with a change in the log Evidence of −4 when using the more complex binary model. Beyond this, however, the evidence favours the CB model, with a change in the log Evidence of over 5000 even when sampling only 3.5 per cent of the orbit. The difference between the models is clear in Fig. 6, where we show a subsection of the real (left-hand panels) and imaginary (right-hand panels) data, zoomed in around the first harmonic of the signal for each of the four simulations. While in the top panel the two models are indistinguishable, differences already become apparent by eye for the 0.1 d orbit, which increase significantly as the orbital period shrinks. In all cases, when using the CB model we recover orbital parameters that are consistent with the injected values.

The purpose of these first simulations was simply to show that, as expected, the acceleration-only model breaks down after a

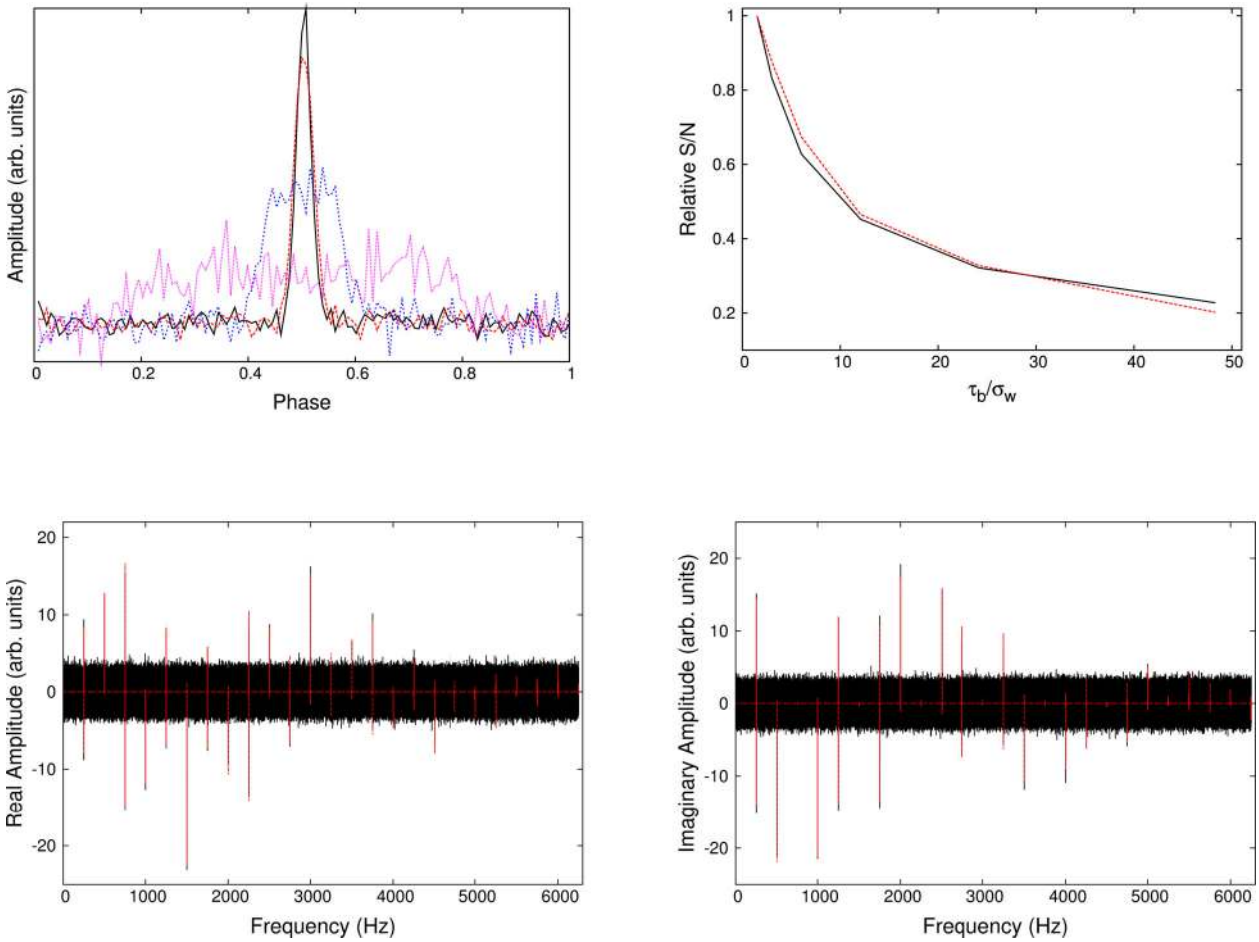

Figure 5. (Top-left) The observed pulse profile when using a sampling time of 80 μs (black lines), 160 μs (red lines), 640 μs (blue lines) and 2560 μs (magenta lines). The FWHM of the intrinsic pulse profile is 120 μs. (Top-right) Predicted S/N loss as a result of the observed pulse width increasing as the sampling time, τ_b increases (black line), and the S/N of the signal extracted by our analysis with COBRA (red line). (Bottom) The real (left) and imaginary (right) part of the model (red lines) and data (black lines) for the 80 μs sampled data in our simulation. In this case, the signal is spread across a large number of harmonics across the sampled Fourier space. Our analysis recovers all the available power in this, and the more extreme cases.

relatively short fraction of the orbit (see e.g. Ransom, Cordes & Eikenberry 2003 for other examples of this in the literature). However, in this instance we are including only a single epoch in our analysis, and thus, as seen in equation (54), we are equally weighting the entire data set. In the following simulation, we will show that, when more optimally weighting the orbit as a function of time, an acceleration model is able to extract an order of magnitude greater S/N compared to the equally weighted scenario.

4.3.2 Eccentric binary

We next simulate one complete orbit of an EB system with parameters as given in Table 1 for PSRA, and show the induced perturbation in the arrival time of the pulses in the top panel of Fig. 7. We generate the simulation using 200 μs sampling, and split the data set up into 10, 12 min epochs, each with 2^{22} samples and a S/N of approximately 20.

As stated previously, the purpose of this simulation is to show that, even with a short period, highly EB, COBRA is able to obtain much more optimized results using simple acceleration-only, or CB models by automatically weighting the different parts of the orbit, compared to equally weighting the data. In particular, this means that faster models can be applied to longer data sets, without being concerned with losing the majority of the signal from smearing out the pulse. While this simulation is still high S/N, we will explore how COBRA behaves with low significance observations in Sections 6 and 7 (cf. also Section 4.1).

In the bottom three panels of Fig. 7, we show the cumulative (black lines), and individual (red lines) S/N of the ten epochs from analyses using (i) an acceleration-only model (top), (ii) the CB model (middle) and (iii) the EB model. In each case, the analysis was performed when including all ten epochs simultaneously. As each epoch has its own overall amplitude parameter (cf. equation 55), this means that the global analysis is able to down-weight parts of the orbit that are a poor match to the model. In this case, as both the simulated instrumental noise, and amplitude of the signal was uniform across the observation, these weights are directly proportional to the individual S/N contributed by each epoch, shown by the red lines.

This is clearly seen in Fig. 7 by comparing the individual S/N that each epoch contributes to the global fit for both the acceleration, and CB models. Both down-weight the epochs near periastron, where the impact of the large eccentricity (0.6) is most apparent. As can be expected, the CB model still fairs better, with a cumulative S/N of approximately 60, compared to 50 for the acceleration-only model. The EB model naturally results in

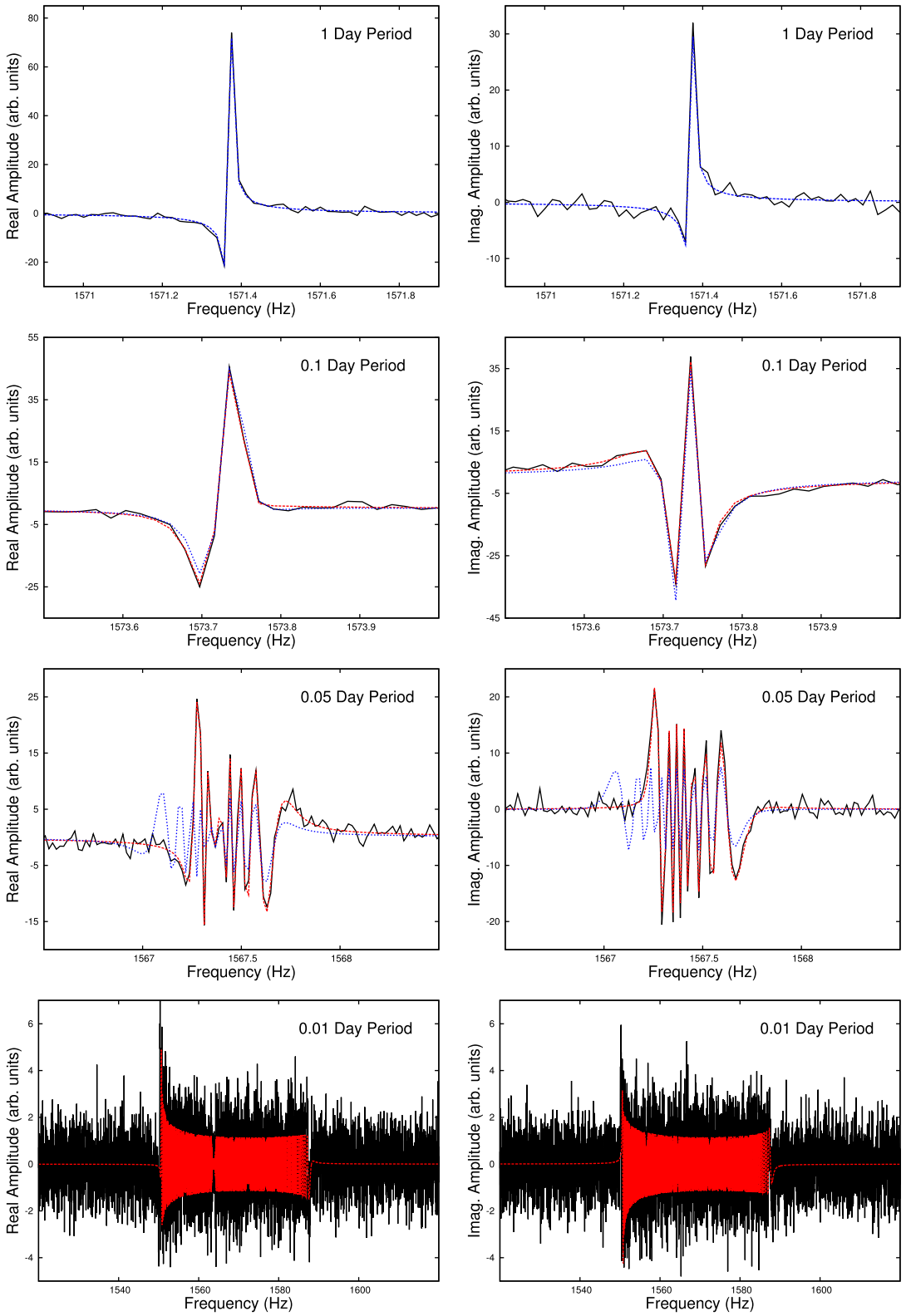

Figure 6. From top to bottom: Simulated 5 min observations of a CB with orbital periods of (i) 1 d, (ii) 0.1 d, (iii) 0.05 d and (iv) 0.01 d. In each case, we show the real (left-hand panels) and imaginary (right-hand panels) components of the data (black lines), and the model when assuming just acceleration (blue lines) and when including a CB (red lines). For the 1 d orbit, both the two models are identical, while for the 0.01 d orbit we plot only the CB model for clarity.

Table 1. Timing model parameters for two simulated eccentric systems.

Name	PSRA	PSRB
Right ascension, α (hh:mm:ss)	00:00:00	00:00:00
Declination, δ (dd:mm:ss)	00:00:00	00:00:00
Pulse period, P (ms)	20.833333	4.35
Orbital period, P_b (d)	0.14286	0.013889
Epoch of periastron, T_0 (MJD)	56000.0	56000.0
Longitude of periastron, ω_0 (deg)	50.0	90.0
Orbital eccentricity, e	0.6	0.1
Projected semimajor axis of orbit, x (s)	0.1	1.6998085
Sine of inclination angle, sin i	–	0.9997
Mass of the pulsar, M_p ($M_\odot$)	–	1.4
Mass of the companion, M_c ($M_\odot$)	–	30
S/N of simulation	70	225

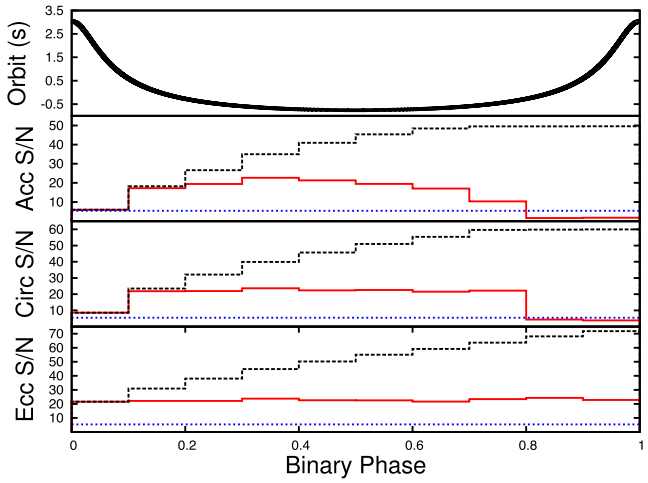

Figure 7. Comparison of the recovered S/N for three different models of the highly eccentric system ($\epsilon = 0.6$) analysed in Section 4.3.2. The change in the pulse arrival times across the simulated orbit as a result of the binary is shown in the top panel. In the following three panels we then show: (i) the S/N contributed per 12 min epoch (red line), (ii) the cumulative S/N (black line) and (iii) the S/N returned by PRESTO from a search of the full 2 h data set (blue line) for the acceleration model (second panel), CB model (third panel) and EB model (bottom panel).

the highest overall S/N, with all epochs contributing equally to the fit, resulting in a total S/N of 70 over the course of the orbit.

For comparison, in each of the three panels we show a blue horizontal line with a value of 5.4, which is the S/N obtained when performing an acceleration search across the full orbit under the assumption of equal weighting. This, in precisely the same way as in Section 4.1, is because after a very short period of time the acceleration model is a sufficiently poor fit that the data starts 'adding noise' to the fit, reducing the S/N that can be achieved from the analysis. By appropriately weighting the different parts of the orbit, however, we can achieve a much greater S/N of 50 for the acceleration-only model, which is only a factor of 1.4 less than the optimal EB model, compared to the over one order of magnitude reduction in S/N achieved from the simpler equally weighted model.

4.3.3 Relativistic binary

In our final simulation, we create an 8 h search mode observation, in which we include a highly relativistic system with parameters given in Table 1 for PSRB. These parameters are chosen such that we maximize the relativistic contributions to the signal. For example, the extreme companion mass (30 $M_\odot$) will result in a significant $\dot{\omega}_b$, and in combination with being almost edge-on to our line of sight will induce a large Shapiro delay. Given the short, 20 min orbital period we have 24 complete orbits in the full 8 h data set. While we acknowledge that such a system will be short lived, with a merger time of approximately 4000 yr, the purpose of this simulation is simply to show that even for such systems our analysis method is capable of fully modelling the data.

In the top panel of Fig. 8, we show the difference between the full RB signal, and a purely Keplerian EB model, after refitting for the binary parameters in order to absorb as much of the difference as possible. Even after refitting for the Keplerian parameters the residuals still cover a period in pulse phase. We indicate with red lines the FWHM of the simulated pulse profile. Any residual outside these lines can be regarded as a near complete loss of signal.

Unlike the previous examples, we approach this simulation as we would a blind search. We initially split the data set up into 512, 1 min epochs, each containing 2^{19} samples, and perform a standard acceleration search using PRESTO. These epochs are chosen to be sufficiently short that an acceleration-only model will still be sufficient to describe the data, and we find that the search recovers multiple candidates at the 7σ confidence level. While this is still relatively large, and would likely be detected in such a short observation using phase-modulation techniques with *PRESTO*, we will explore lower S/N data in Section 7. We choose the most significant of these candidates, and use the period returned to constrain the prior in our analysis, with a width of ± 1 per cent around that value.

In the second panel of Fig. 8, we show how the S/N grows as we include more 1 min epochs into the analysis, and progressively increase the complexity of the model used. For example, even just using a single 1 min epoch, the CB model is able to better describe the data. We find the difference in the log Evidence for the CB model compared to the acceleration-only model is 22 for this first epoch, and correspondingly, the S/N obtained increases from 7.5σ to 10σ.

As we incorporate more data, and the precision of the parameter estimates we obtain improves, we can use the posteriors from the current analysis to update the size of the priors we use in the subsequent stage of analysis that includes more data. As such, even though the parameter space for the full relativistic model may seen extreme in the context of the blind search, by the time the EB model is no longer sufficient to model the data appropriately, we have already significantly reduced our priors on the Keplerian parameters, and thus only need broad priors for the pulsar mass of [0.5, 3 $M_\odot$], and the companion mass, of [0.1, 30 $M_\odot$]. As an example, for our initial analysis of the first 1 min epoch, our prior on the binary period covered a range of 10^{-3} up to 1 d. When including two epochs we could already set an upper limit on the binary period of 0.1 d, and after ten epochs our prior only covered a range from 0.005 to 0.02 d.

We find that the EB model is preferred by the data after approximately 10 epochs, corresponding to half an orbit, and that after only one complete orbit the RB model is already preferred, owing to the significant contribution to the binary signal from Shapiro delay, which is clearly visible in the top panel of Fig. 8 as the periodic spikes in the residuals.

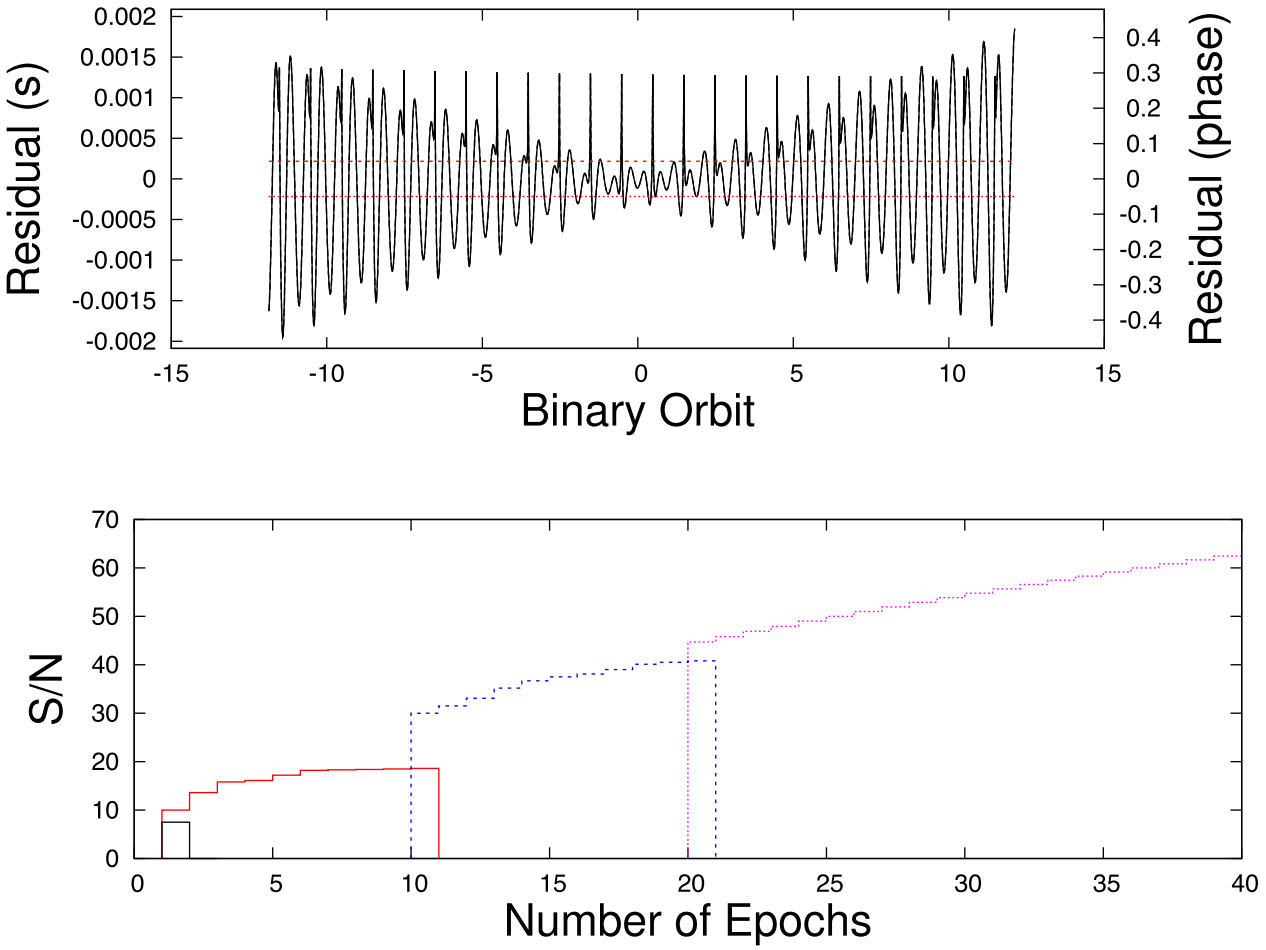

Figure 8. (Top) The difference between the signal induced by the RB system in Table 1 for PSRB, and a purely Keplerian system, after refitting for the binary parameters in order to minimize the difference between the two. Red lines indicate the FWHM in pulse phase of the injected system. Any residual outside these lines can be regarded as a near complete loss of signal. (Bottom) The S/N of our analysis of the RB system as more 1 min epochs are included, and progressively more complex models are used. These models included acceleration-only (black lines), a CB (red lines), an EB (blue lines) and the full relativistic model (magenta lines).

In Fig. 9, we show the one- and two-dimensional posterior probability distributions for the pulsar and companion masses from our analysis of the full 8 h data set using the RB model, as well as the derived posterior probability distributions for three of the PK parameters, $\dot{\omega}_{\rm b}$, $\sin i$ and γ (black lines). The extreme nature of this system is evident by comparing the magnitude of these measurements to the most RB pulsar system currently known (Kramer et al. 2006). In this case, $\dot{\omega}_{\rm b} = 2490 \pm 8\,{\rm deg\ yr^{-1}}$, compared to $16.9\,{\rm deg\ yr^{-1}}$ for the double pulsar. This implies that the orbit has precessed 2.3 deg over the course of the 8 h observation, and so it is not surprising that the relativistic timing model was preferred over the Keplerian model within a single orbit. In the same figure, we also plot the posterior probability distributions for an analysis using the PKB model, where we included the Keplerian parameters directly as free parameters in the analysis (red lines). Both sets of analyses result in consistent parameter estimates.

5 OPTIMIZATION OF EPOCH SIZE

Given our ability to combine different epochs coherently, we can check to see at what stage the computing cost of calling the GPU kernels to evaluate the signal model for each epoch overtakes gains made in performing smaller Fourier transforms. In Table 2, we list the run times in seconds for 10 000 calls of the likelihood using models for an isolated pulsar (Iso), accelerated pulsar (Acc), CB, EB, and finally the RB model, as we split an observation with a total of 2^{25} samples into a larger number of separate epochs.

We find that for a Kepler K40 GPU the optimal run time is for epochs with $2^{21} \rightarrow 2^{22}$ samples. For a sample length of 64 μs, as is typical for most modern surveys, this corresponds to approximately 5 min of data. We note that these exact values will depend both on the GPU architecture used to run the search code, and the version of the code used, as optimizations will result in more time being spent performing FFTs as opposed to building the signal model. Note that there is a second factor that impacts the run time, which is the number of samples required to perform the analysis. The number of samples required by *MultiNest* scales roughly exponentially as the number of dimensions increases, and so while it is extremely efficient for low-dimensional problems (<10) it rapidly becomes a limiting factor for more complex problems. The more recent *PolyChord* (Handley, Hobson & Lasenby 2015) scales as roughly d^3, with d the number of dimensions which may be preferable for the RB model. Further advances in sampling technology will naturally improve this aspect of the problem.

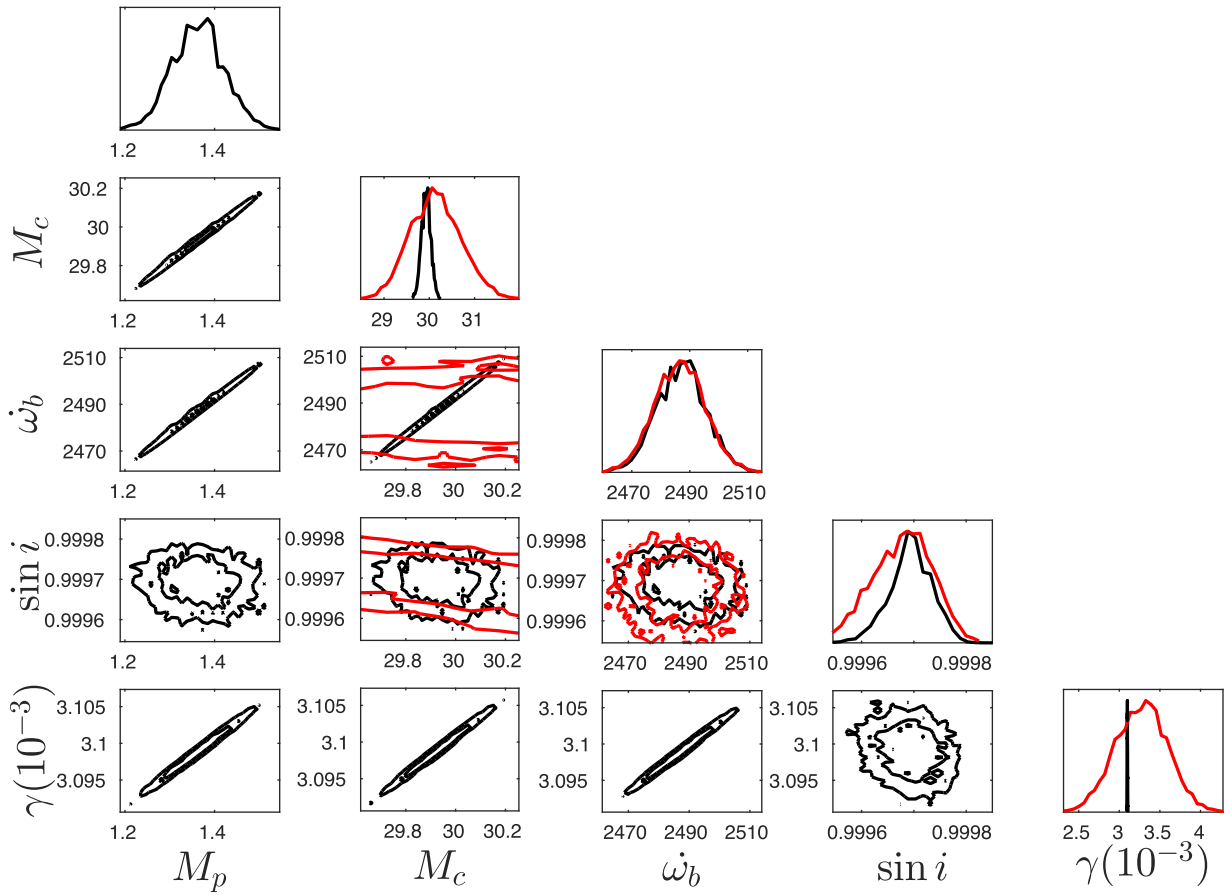

Figure 9. One- and two-dimensional posterior distributions from our analysis of the RB system in Section 4.3.3 when including the full 8 h of search data, using the RB model (black lines) and PKB model (red lines). We show the posterior distributions for the two system masses, and for the three PK parameters that were detected at greater than 3σ significance in the PKB model. For the RB model, the posterior distributions on the PK parameters have been derived directly from the posterior distributions for the two masses.

Table 2. Run times in seconds for 10 000 likelihood evaluations for the different pulsar models described in Section 3.

No. of epochs	Epoch size	Iso	Acc	CB	EB	RB
1	2^{25}	360	360	440	480	520
2	2^{24}	350	350	440	480	520
4	2^{23}	340	340	430	470	510
8	2^{22}	330	330	420	460	500
16	2^{21}	350	350	440	480	470
32	2^{20}	400	400	490	540	580
64	2^{19}	490	490	600	650	700

6 APPLICATION TO OBSERVATIONS OF THE GLOBULAR CLUSTER 47Tuc

We now apply the techniques described in Sections 3 and 4 to a series of observations of the GC 47Tuc, taken with the 64-m Parkes Telescope at a central frequency of 1.4 GHz, with a bandwidth of 256 MHz. The spread in Dispersion Measure (DM) for all known pulsars in the cluster is less than 1 cm^{-3} pc, and so for the purposes of this analysis, we simply dedisperse all the data to a reference DM of 24.6 cm^{-3} pc, which is the DM of the pulsar J0023−7204C. All the data analysed in this section is publicly available on the CSIRO data archive.[5]

[5] https://data.csiro.au/dap/public/atnf/pulsarSearch.zul

In particular, we use COBRA to analyse observations that contain three known pulsars in 47Tuc, the isolated pulsar PSR J0024−7204D (47TucD, Manchester et al. 1990), and the binary pulsars PSR J0023−7203J (47TucJ, Manchester et al. 1991), and PSR J0024−7204R (47TucR, Camilo et al. 2000, henceforth C2000). A summary of the timing parameters for these pulsars are given in Table 3.

First, we use 47TucD to demonstrate the ability of our method to coherently combine sparsely sampled data, where observations have been made over a 100 d period, each separated by weeks. In particular, we include 22 min of data from an observation at MJD 52526.8, and 11 min of data from observations at MJDs 52571.4 and 52630.3. In each case, the data is sampled at a rate of 160μs, and we include epochs containing 2^{19} samples each, which results in 16 epochs from the first day, and eight epochs for the second 2 d of observing. In Fig. 10 (top-left panel), we show the cumulative S/N as we include more epochs (black line) as well as the S/N of each individual epoch (red line). As this is one of the brighter pulsars in 47Tuc the detections are consistently significant (6–7σ) even in these short, 80 s epochs, and so the cumulative S/N grows smoothly as we include more data. In the top-right panel of Fig. 10, we show the one- and two-dimensional posterior parameter estimates for the period and period-derivative from our analysis of the full data set. Due to the extremely sparse coverage of data there are a significant number of equally probable solutions, each shown as an island in the two-dimensional plot. While navigating such

Table 3. Timing model parameters for PSRs J0024−7204D, J0023−7203J and J0024−7204R.

	Parameters from long-term timing		
	J0024−7204D	J0023−7203J	J0024−7204R
Right ascension, α (hh:mm:ss)	00:24:13.8776	00:23:59.4040	00:24:05.67
Declination, δ (dd:mm:ss)	−72:04:43.8323	−72:03:58.7720	−72:04:52.62
Pulse period, P (ms)	5.3575732850382	2.1006335458586	3.480463
First derivative of pulse period, $\dot{P}$ (10^{-20})	−0.333	−0.9787	–
Orbital period, P_b (d)	–	0.1206649386	0.0662
Orbital eccentricity, e	–	<0.0003	–
Projected semimajor axis of orbit, x (s)	–	0.0404087	0.0334

a parameter space is challenging for standard MCMC methods, MULTINEST[6] excels in such areas (cf. the 'eggbox' problem in Feroz, Hobson & Bridges 2009). As in Section 4.3.3, we decrease the size of the priors on the parameters included in the model as we include more data into the analysis. As such, being able to explore the full range of solutions at every stage is critical to ensure that the global maximum remains in our prior.

The two binary pulsars chosen, 47TucJ and 47TucR, have the two shortest binary periods of any known pulsar in 47Tuc, of 2.9 and 1.6 h, respectively. In our analysis, we use 1 h of data for 47TucJ taken at MJD 52572.6, and 2 h of data for 47TucR taken at 52462.0, and so will be sampling approximately one-third of an orbit, and over one complete orbit respectively for these two pulsars. We split the data into 12 min epochs of 2^{22}, 160 μs samples, for both pulsars, and as for 47TucD, in Fig. 10 (left-middle, left bottom panels) we plot the cumulative S/N achieved as we perform our analysis using the CB model, while including an increasing number of blocks (black lines) along with the S/N of each 12 min epoch individually (red lines). While 47TucJ is consistently bright across the included observations, 47TucR provides a more robust test of our analysis method, as it can be seen to dip significantly in brightness in epochs six and seven. Comparing this with the cumulative S/N, we can see that when these epochs are included in the analysis they do not lead to a decrease in the cumulative S/N, as would be the case if they were combined under the assumption of equal weighting. Rather, as in Section 4.1, by weighting these noisy epochs appropriately in the global fit, the cumulative S/N simply flattens across those epochs as they do not contribute significantly to the analysis.

In the right-panels of Fig. 10, we show the one- and two-dimensional marginalized posterior distributions for the binary amplitude, binary phase and binary period, as the number of epochs in the analysis is increased. For 47TucJ we include results from two, three and five epochs (middle), and for 47TucR we show the results for three, six and ten epochs.

In all three cases, the parameter estimates are consistent with values determined from long-term timing regardless of the number of epochs included in the analysis. For 47TucR in particular we achieve higher precision than that published in C2000, where, for example, we find the binary period to be 0.066260(12) d compared to 0.0662(1) d in C2000. We stress, however, that in neither case has there been detailed modelling of the noise in the search mode observations, and so it is not clear to what extent this will bias the parameter estimates and their uncertainties. Despite this, it is still of interest to test more complex models, and in particular we perform a model comparison between the circular, and EB models. We obtain a difference in the log Evidence of −1 when using the EB model, indicating the data does not support the use of the more complex model. This can be seen in the parameter estimates for eccentricity, which we show in Fig. 11 for both a prior that is uniform in the log of the eccentricity (black lines) and uniform in the eccentricity (red lines). In both cases, we simply recover the prior, and set a 95 per cent upper limit on the eccentricity of 0.0007, indicated by the red line in Fig. 11.

Finally, we use 47TucR to explore the impact of using a simple Gaussian model for the pulse profile, given that in this case the profile has two clear peaks (cf. C2000 Fig. 5). In principle, the maximum loss in S/N that we could expect is 50 per cent for the case where both peaks are equal in amplitude. We use the PSRCHIVE tool PADD to combine the ten epochs used in our analysis, and compute the S/N loss due to the mismatch from our Gaussian template. We find a 40 per cent loss, consistent with our expectation given the relative area of the two peaks. Given this loss, a more optimal search would include the option for more complex models for the pulse profile, either a double Gaussian which would increase the dimensionality by three, or a more general shapelet model (Refregier 2003), which could result in an arbitrary increase in dimensionality depending on the complexity of pulse shape desired. Such extensions will be considered in the future, and will only become more tractable as sampling technology improves.

7 USING PHASE INFORMATION TO DISCRIMINATE AGAINST RFI

In Section 2, we discussed how, in Bayesian analysis, the significance of a potential source can be determined by comparing the Evidence for two different models (such as a pulsar being present in the data, compared to no pulsar).

As for Frequentist statistics, this significance depends heavily on the noise model used in the analysis. If there is significant RFI present in the data set, which is not incorporated into the model for the noise, then the significance of a source will likely be overstated. Clearly, however, modelling RFI is not trivial, and one would like to have an objective means of determining the significance of a candidate source, independent of how well one can model the noise.

In this section, we suggest one such approach towards achieving this goal, which leverages the fact that COBRA does not just account for the power in the signal, but fully incorporates phase information into the analysis. It should therefore be possible to use this phase information to discriminate between a pulsar, that is described by a physical timing model, and a source of RFI, as long as that source of RFI has less phase information over the observed time-scales than the pulsar. Even if a source of RFI is pulsed at regular intervals, by combining search data taken over many months the combination of barycentering the data, and any additional corrections required

[6] https://ccpforge.cse.rl.ac.uk/gf/project/multinest

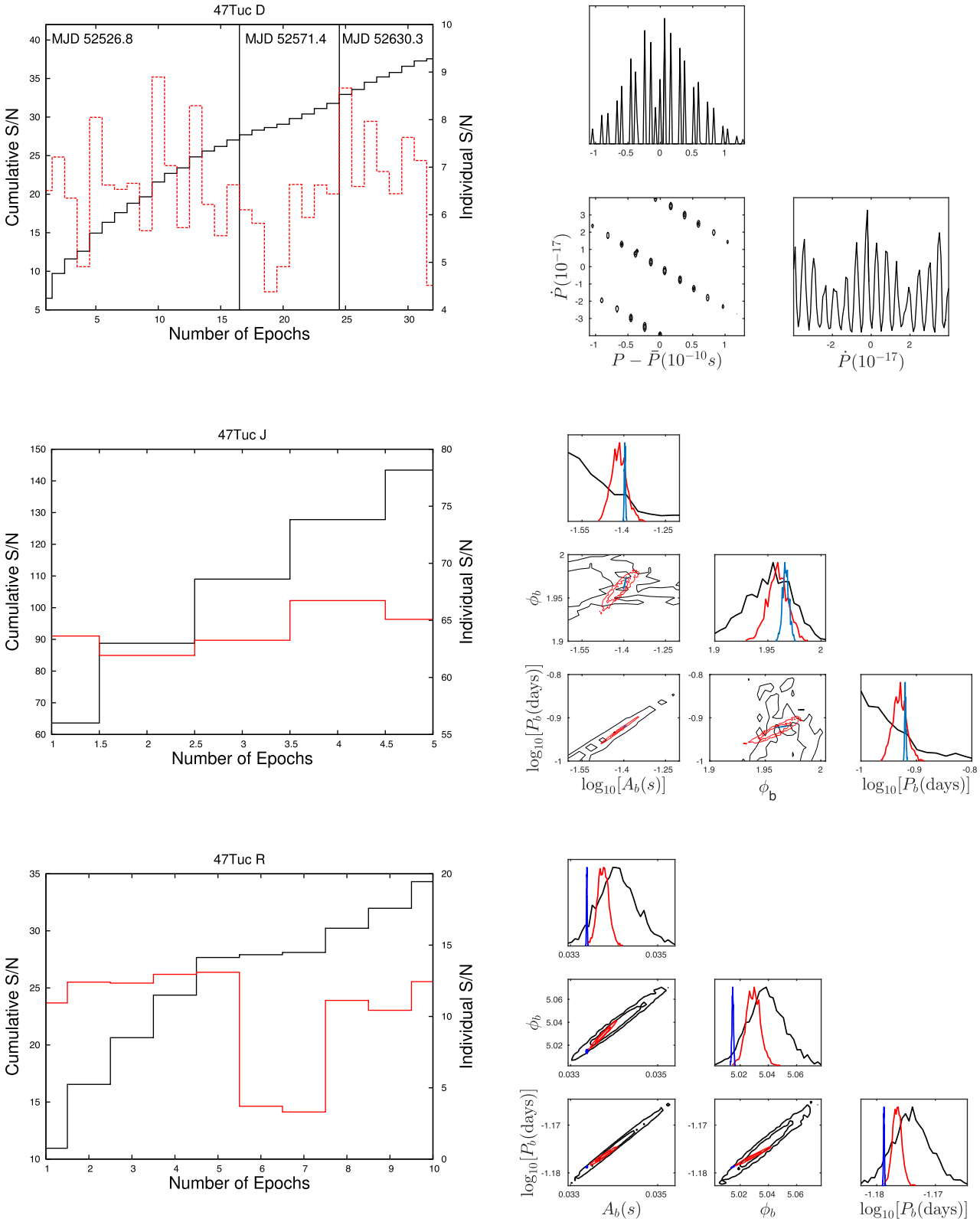

Figure 10. (Left-hand panels) Cumulative S/N (black lines, left *y*-axis) as additional epochs are added into the analysis for PSRs J0024−7204D (top), J0023−7203J (middle) and J0024−7204R (bottom). In each case, we compare this to the S/N obtained from each epoch (red lines, right *y*-axis). (Right-hand panels) One- and two-dimensional marginalized posterior distributions for (i) the period and period derivative of PSR J0024−7204D from an analysis of the full data set (top), (ii) the binary amplitude, binary phase, and binary period of PSR J0023−7203J (middle) for two (black lines), three (red lines) and five (blue lines) epochs and (iii) the same binary parameters for PSR J0024−7204R for three (black lines), six (red lines) and ten (blue lines) epochs.

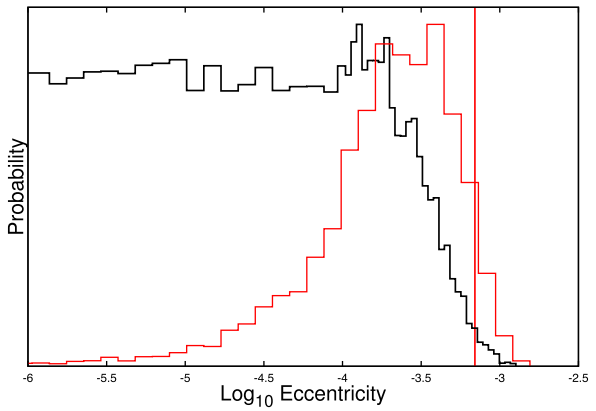

Figure 11. One-dimensional marginalized posterior probability distribution for $\log_{10}\epsilon_b$ from our analysis of 47TucR using a prior that is uniform in the log of the eccentricity (black line) and uniform in the eccentricity (red line). No significant detection of eccentricity is made, with our analysis returning the prior in both cases. Using the prior that is uniform in the eccentricity, we set a 95 per cent upper limit of 0.0007 indicated by the vertical red line.

by the timing model either due to a binary companion or intrinsic spin-down, will enable these signals to be discriminated from one another.

Our goal is therefore to modify the data set in such a way that no real pulsar signal should persist, but to keep the statistical properties of the RFI and other sources of noise intact. We can then use the Evidence for a pulsar in this modified data set to 'calibrate' our significance. This kind of calibration has been performed in the context of gravitational wave analysis using a pulsar timing array (Taylor et al. 2016) and we apply it to our search method for new pulsars.

In order to modify our data set, we simply multiply each Fourier frequency by a complex number $\exp(i\theta)$, where θ takes a random value from zero to 2π. Each frequency is assigned a different random number, and this has the effect of 'scrambling' the phase of the signal, without affecting the power. Thus, a reanalysis of this modified data set with COBRA should results in the detection of only noise-like signals.

We demonstrate the efficacy of this approach using two examples, one using the pulsar 47TucD described previously, and second using a source of RFI.

In Fig. 12 (left-hand panel), as for the data sets in Section 6, we show the cumulative S/N as additional epochs are added into an analysis of PSR 47TucD, and compare this to the S/N obtained from each epoch individually. In this case, we have used only the first eight epochs from MJD 52526, and have split the last seven epochs up into blocks of four, so that the S/N contributed by each one is decreased by approximately a factor of 2. Thus, while the first epoch provides a 'high' S/N of 6.7σ, the remaining 28 epochs have a S/N of only $3 \pm 1\sigma$, and so would likely not be considered significant, especially in light of the potentially poorly modelled noise estimate that has been used to determine these ratios.

In Fig. 12 (middle panel), we show a histogram for the change in the log Evidence between a model for an isolated pulsar, compared to no pulsar, for 50 independent phase scrambles of this data set, where each epoch used in the analysis has also been scrambled independently. These scrambles result in a distribution of Δlog Evidences which we then shift so that the mean is at zero. We can then compare this with the Δlog Evidence for the non-scrambled data set, shifted by the same amount, for which we obtained a value of 124 (vertical red line in the panel). This value is therefore completely inconsistent with the distribution of 'noise-like' Evidences, indicating that there is significant phase information contributing to the signal for this pulsar, as we would expect for a real source. Thus, even when individual epochs contribute relatively little signal, by phase coherently combining these epochs we can still obtain a significant detection.

We contrast this with the right-hand panel of Fig. 12, which shows the equivalent histogram for a candidate pulsar at a frequency of 34.6 Hz. In this case, the Δlog Evidence of the candidate in the non-scrambled data set was 430, which, if we believed our noise model, would imply an extremely significant detection of a new pulsar.

If we perform 50 phase scrambles of the data set, however, we find a distribution of Δlog Evidences equal to 344 ± 46. Shifting both values by the mean of this distribution, we see that the candidate falls within the histogram of noise-like data sets, and thus we discard the candidate as insignificant.

8 CONCLUSIONS

We have introduced COBRA, a Bayesian toolkit that can perform the roles of both pulsar searching and pulsar timing on single pulse time series data. Rather than work with the power spectrum, as is common for pulsar search techniques, COBRA constructs a model

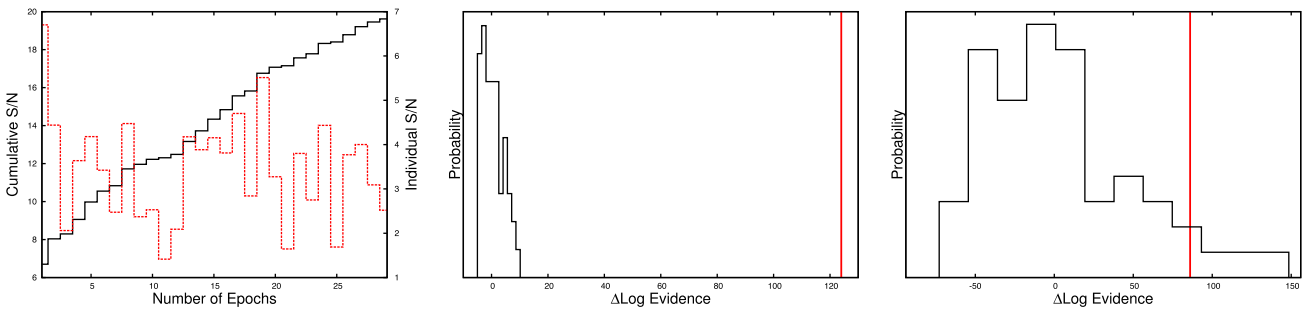

Figure 12. (Left-hand panel) Cumulative S/N (black lines, left y-axis) as additional epochs are added into an analysis of PSR 47TucD. In each case, we compare this to the S/N obtained from each epoch (red lines, right y-axis). Compared to the data set used in Section 6 we have split the epochs into smaller blocks so that the S/N contributed by each one is less, and only used the first eight epochs from MJD 52526. (Middle and right-hand panels) Histograms of the change in log Evidence for a model that includes a pulsar, compared to a model that does not for 50 phase scrambles of the 47TucD data set (middle panel), and a source of RFI (right-hand panel). In both cases, we indicate the change in log Evidence for the same models in the unscrambled data sets as a vertical red line.

pulse train that can incorporate effects such as aliasing, scattering and binary motion, and compares this directly to the data. This means that COBRA performs a fully phase coherent timing analysis, allowing it to combine data over multiple epochs, frequencies, or even telescopes.

We showed that our analysis methods can deal with the effects of scintillation, where the brightness of the pulsar can dip by orders of magnitude, by incorporating the amplitude of the signal at different epochs as free parameters in the analysis. These amplitudes are then marginalized over analytically to avoid increasing the dimensionality of the parameter space that must be searched over numerically. This provides an added benefit when fitting a model to the data that is not sufficient to fully describe the system, for example, when using a CB model to describe a highly eccentric system. In such an analysis, the parts of the orbit that do not match the model will be down weighted, ensuring that the maximum S/N ratio will be extracted from the data.

We applied COBRA to a series of simulations that included aliased signals, and highly relativistic binaries, using candidates from standard acceleration searches to limit the prior range on the source periodicity, and performing a blind search in any remaining parameters in our Bayesian analysis. Even in an 8 h data set that contained 24 full orbits of a pulsar in a binary with a 30 $M_\odot$ companion, we showed that this approach allows us to build up a model that fully describes the signal, including PK effects.

Finally, we then analysed observations of the GC 47Tuc that contained the isolated pulsar, PSR J0024−7204D (47TucD), and the binary pulsars PSR J0023−7203J (47TucJ) and PSR J0024−7204R (47TucR). For 47TucD we performed the search coherently across observations made over a period of 100 d, where each observation was separated by weeks, while for the binary pulsars, we included 1 h of data for 47TucJ, and 2 h of data for 47TucR, which covered one-third of an orbit, and over one complete orbit, respectively. As with the simulations, we were able to go from an initial search candidate, to obtaining full timing solutions from the search data alone. For 47TucR, we also compared circular and EB models, however found no evidence for eccentricity in the data analysed, setting a 95 per cent upper limit of $\epsilon_b < 0.0007$.

The recent discovery of merging stellar mass black holes (Abbott et al. 2016) proved the existence of such extreme objects, but so far no evidence has been found for a pulsar orbiting either a stellar mass black hole, or the supermassive black hole in the centre of our Galaxy. However, extremely long search baselines for both GCs and the Galactic Centre have been established over the last decade, and an optimal, coherent search of such data may provide a means for proving the existence of such a system. In addition, new endeavours to search for pulsars in extreme settings are going to come online in the near future, such as the BlackHoleCam project (Goddi et al. 2017). Advanced search algorithms like COBRA will be key to fully exploiting the data obtained from these projects, and maximizing the probability of detecting groundbreaking new systems.

ACKNOWLEDGEMENTS

The Parkes radio telescope is part of the Australia Telescope National Facility which is funded by the Commonwealth of Australia for operation as a National Facility managed by Commonwealth Science and Industrial Research Organization (CSIRO). We thank all of the observers, engineers and Parkes observatory staff members who have assisted with the observations reported in this paper. LL was supported by a Junior Research Fellowship at Trinity Hall College, Cambridge University. PT gratefully acknowledges financial support by the European Research Council for the ERC Starting grant BEACON under contract No. 279702. PT is supported for this research through a stipend from the International Max Planck Research School (IMPRS).

REFERENCES

Abbott B. P. et al., 2016, Phys. Rev. Lett., 116, 061102
Abdo A. A. et al., 2009, Science, 325, 840
Abdo A. A. et al., 2010, A&A, 524, A75
Aulbert C., 2007, preprint (arXiv: astro-ph/0701097)
Barr E. D. et al., 2013, MNRAS, 435, 2234
Camilo F., Lorimer D. R., Freire P., Lyne A. G., Manchester R. N., 2000, ApJ, 535, 975
Champion D. J. et al., 2010, ApJ, 720, L201
Coenen T. et al., 2014, A&A, 570, A60
Cordes J. M. et al., 2006, ApJ, 637, 446
Damour T., Deruelle N., 1986, Ann. Inst. Henri Poincaré Phys. Théor., 44, 263
Damour T., Taylor J. H., 1991, ApJ, 366, 501
Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601
Goddi C. et al., 2017, Int. J. Mod. Phys. D, 26, 1730001
Handley W. J., Hobson M. P., Lasenby A. N., 2015, preprint (arXiv e-prints)
Hewish A., Bell S. J., Pilkington J. D. H., Scott P. F., Collins R. A., 1968, Nature, 217, 709
Kass R. E., Raftery A. E., 1995, J. Am. Stat. Assoc., 90, 791
Keith M. J. et al., 2010, MNRAS, 409, 619
Keith M. J. et al., 2013, MNRAS, 429, 2161
Kramer M. et al., 2006, Science, 314, 97
Liu K., Eatough R. P., Wex N., Kramer M., 2014, MNRAS, 445, 3115
Lorimer D. R., 2011, Astrophysics Source Code Library, record ascl:1107.016
Lorimer D. R., Kramer M., 2005, Handbook of Pulsar Astronomy. Cambridge Univ. Press, Cambridge
Malofeev V. M., Malov O. I., Shchegoleva N. V., 2000, Astron. Rep., 44, 436
Manchester R. N., Lyne A. G., Johnston S., D'Amico N., Lim J., Kniffen D. A., 1990, Nature, 345, 598
Manchester R. N., Lyne A. G., Robinson C., Bailes M., D'Amico N., 1991, Nature, 352, 219
Maron O., Kijak J., Kramer M., Wielebinski R., 2000, A&A, 147, 195
Pan Z., Hobbs G., Li D., Ridolfi A., Wang P., Freire P., 2016, MNRAS, 459, L26
Ransom S., 2011, Astrophysics Source Code Library, record ascl:1107.017
Ransom S. M., Cordes J. M., Eikenberry S. S., 2003, ApJ, 589, 911
Ransom S. M., Hessels J. W. T., Stairs I. H., Freire P. C. C., Camilo F., Kaspi V. M., Kaplan D. L., 2005, Science, 307, 892
Refregier A., 2003, MNRAS, 338, 35
Sivia D., Skilling J., 2006, Data Analysis: A Bayesian Tutorial. Oxford Science Publications
Smith K. M., 2016, preprint (arXiv e-prints)
Smits R., Kramer M., Stappers B., Lorimer D. R., Cordes J., Faulkner A., 2009, A&A, 493, 1161
Soglasnov V. A., Popov M. V., Bartel N., Cannon W., Novikov A. Y., Kondratiev V. I., Altunin V. I., 2004, ApJ, 616, 439
Sulman B., Ransom S., Stinebring D., 2005, Bull. Am. Astron. Soc., 37, 1469
Tauris T. M., Savonije G. J., 1999, A&A, 350, 928
Taylor J. H., Weisberg J. M., 1989, ApJ, 345, 434
Taylor S. R., Lentati L., Babak S., Brem P., Gair J. R., Sesana A., Vecchio A., 2016, Phys. Rev. D. 95, 042002
Turk P. J., Lorimer D. R., 2013, MNRAS, 436, 3720
Weisberg J. M., Huang Y., 2016, ApJ, 829, 55
Williamson I. P., 1972, MNRAS, 157, 55
Wolszczan A., Frail D. A., 1992, Nature, 355, 145

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