

# Kalman tracking and estimation of continuous gravitational waves with a pulsar timing array: inclusion of the pulsar terms

Tom Kimpson<sup>1,2\*</sup>, Andrew Melatos<sup>1,2</sup>, Joseph O’Leary<sup>1,2</sup>, Julian B. Carlin<sup>1,2</sup>, Robin J. Evans<sup>3</sup>, William Moran<sup>3</sup>, Tong Cheunchitra<sup>1,2</sup>, Wenhao Dong<sup>1,2</sup>, Liam Dunn<sup>1,2</sup>, Julian Greentree<sup>3</sup>, Nicholas J. O’Neill<sup>1,2</sup>, Sofia Suvorova<sup>3</sup>, Kok Hong Thong<sup>1,2</sup>, Andrés F. Vargas<sup>1,2</sup>

<sup>1</sup>*School of Physics, University of Melbourne, Parkville, VIC 3010, Australia*

<sup>2</sup>*OzGrav, University of Melbourne, Parkville, VIC 3010, Australia*

<sup>3</sup>*Department of Electrical and Electronic Engineering, University of Melbourne, Parkville, Victoria 3010, Australia*

Last updated 6 October 2023

## ABSTRACT

The detection of continuous nanohertz gravitational waves from individual supermassive black hole binaries can be formulated within a state-space framework, wherein intrinsic achromatic spin wandering is tracked simultaneously with the modulation induced by a single-source gravitational wave in the pulse times of arrival. Such an approach complements traditional analysis methods by enabling self-consistent inference of model parameters and calculation of the Bayesian evidence, without fitting for ensemble-averaged statistics such as the power spectral density of the timing noise. Previous analyses have worked exclusively with the ‘Earth term’, leading to biases in the parameter estimation. Here it is shown how to extend previous analyses to include the so-called ‘pulsar terms’. It is shown via astrophysically representative software injections in Gaussian measurement noise that the inclusion of the pulsar terms increases the parameter accuracy relative to the Earth terms approach by  $X$  and the sensitivity by  $Y$ .

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

## 1 INTRODUCTION

Nanohertz gravitational waves (GWs) can be produced by the inspiral of supermassive black hole binaries (SMBHBs; Rajagopal & Romani 1995; Jaffe & Backer 2003; Wyithe & Loeb 2003; Sesana 2013; McWilliams et al. 2014; Ravi et al. 2015; Burke-Spolaor et al. 2019; Sykes et al. 2022). A nHz GW influences the times of arrival (TOAs) of radio pulses from pulsars measured at the Earth. A pulsar timing array (PTA Tiburzi 2018; Verbiest et al. 2021) simultaneously measures the TOAs from multiple pulsars. In this way it is possible to detect GWs at frequencies that are prohibitively low for terrestrial interferometers.

PTAs search for two types of SMBHB GW signals. The first is a stochastic GW background due to the incoherent superposition of multiple SMBHB sources (Allen 1997; Sesana et al. 2008; Christensen 2019; Renzini et al. 2022). Consilient evidence for the GW background has been presented by multiple separate PTAs (Agazie et al. 2023a; Antoniadis et al. 2023a; Reardon et al. 2023; Xu et al. 2023), by measuring the cross-correlation between pairs of pulsars in the pulsar timing residuals as a function of the angular separation between the pulsars – the Hellings-Downs curve (Hellings & Downs 1983). The second is resolved GWs from sufficiently close and massive individual SMBHBs (Sesana & Vecchio 2010; Yardley et al. 2010; Zhu et al. 2015a; Babak & Sesana 2012a; Ellis 2013;

Zhu et al. 2016). Individual SMBHBs are sources of continuous GWs. This results in persistent, quasi-monochromatic modulations in the pulsar timing residuals which can be modelled. Therefore, rather than cross-correlating pulsar pairs, individual SMBHBs are detected by either frequentist matched filtering (Lee et al. 2011; Ellis et al. 2012; Zhu et al. 2014) or Bayesian inference (Ellis & Cornish 2016; Arzoumanian et al. 2020a). Observational efforts to detect individual sources have been unsuccessful so far (Jenet et al. 2004; Zhu et al. 2014; Babak et al. 2016; Arzoumanian et al. 2023), although some inconclusive evidence at low significance has been presented (Antoniadis et al. 2023c).

Irregularities in the intrinsic rotation of the pulsar result in random, unmodelled, red-spectrum TOA fluctuations known as ‘timing noise’. Timing noise limits the sensitivity of PTAs to GW signals (Shannon & Cordes 2010a; Lasky et al. 2015; Caballero et al. 2016; Goncharov et al. 2021). Additional white noise sources such as phase jitter noise and radiometer noise (Cordes & Shannon 2010; Lam et al. 2019; Parthasarathy et al. 2021) must also be considered. Kimpson (2023) (K23 hereafter) presented an alternative and complementary approach to PTA data analysis for individual, quasi-monochromatic, SMBHB sources which self-consistently tracks the intrinsic timing noise in PTA pulsars and disentangles it from GW-induced TOA modulations. Specifically, K23 formulate PTA analysis as a state-space problem and use a Kalman filter (Kalman 1960) to track the pulsars’ intrinsic rotational state in conjunction with a Bayesian nested sampler (Skilling 2006; Ashton

\* Contact e-mail: [tom.kimpson@unimelb.edu.au](mailto:tom.kimpson@unimelb.edu.au)

et al. 2022) for parameter estimation and model selection.

It is common in standard PTA analysis to drop the so called "pulsar term" from the modelling (e.g. Sesana & Vecchio 2010; Babak & Sesana 2012a; Petiteau et al. 2013; Zhu et al. 2015b; Taylor et al. 2016; Goldstein et al. 2018; Charisi et al. 2023). Unlike the "Earth term", the pulsar term is not phase coherent and cannot be summed across the array to increase the total signal-to-noise. Instead it is considered as a source of self noise. However, dropping the pulsar terms results in a reduced detection probability ( $\sim 5\%$ ) bias in the parameter estimates (Zhu et al. 2016; Chen & Wang 2022). The method presented in K23 also drops the pulsar term. In this paper we remove this limitation and present the state-space approach to PTA data analysis including the pulsar term.

This paper is organised as follows. In Section 2 we briefly review the state-space formulation of PTA data analysis introduced in Kimpson (2023). We do not review the methods for state tracking, parameter inference and model selection and refer the reader to K23. In Section 3 we introduce how to include the pulsar term in the state-space model. In Section 4 we test the updated method on the synthetic data for a single representative example GW system, and compare the results to the original Earth-terms model. In Section 8 we extend the tests to cover an astrophysically relevant parameter space. Discussion on future extensions and conclusions are made in Sections 10 and 11 respectively. Throughout the paper we adopt the natural units, with  $c = G = \hbar = 1$ , and metric signature  $(-, +, +, +)$ .

## 2 STATE-SPACE FORMULATION

In this section we briefly review the state-space formulation of PTA analysis presented in (Kimpson 2023).

The intrinsic state variable is the  $n$ -th pulsar's spin frequency  $f_p^{(n)}(t)$ , as measured in the local, freely-falling rest frame of the pulsar's centre of mass. The superscript  $1 \leq n \leq N$  indexes the  $n$ -th pulsar in the array.  $f_p^{(n)}(t)$  evolves according to a stochastic differential equation and is related to the radio pulse frequency measured by an observer at Earth,  $f_m^{(n)}(t)$  via a measurement equation.

A phenomenological model for the evolution of  $f_p^{(n)}(t)$  is presented in Section 2.1. The measurement equation relating  $f_m^{(n)}(t)$  to  $f_p^{(n)}(t)$  is presented in Section 2.2.

### 2.1 Spin evolution

In our model

Equation X is inherently an idealized phenomenological model to capture the main qualitative features of a typical PTA pulsar's observed spin evolution, i.e. random, small-amplitude excursions around a smooth, secular trend. At present, a predictive, first-principles theory of timing noise does not exist and there are several plausible physical mechanisms, referenced in Section 1.

In the model,  $f_p^{(n)}(t)$  evolves according to the sum of a deterministic torque and a stochastic torque. The deterministic torque is attributed to magnetic dipole braking, with braking index  $n_{\text{em}} = 3$  (Goldreich & Julian 1969). Most PTAs involve millisecond pulsars, for which the quadratic correction due to  $n_{\text{em}}$  in  $f_p^{(n)}(t)$  is negligible over the observation time  $T_{\text{obs}} \sim 10$  yr, and the deterministic

evolution  $f_{\text{em}}^{(n)}(t)$  can be approximated accurately by

$$f_{\text{em}}^{(n)}(t) = f_{\text{em}}^{(n)}(t_1) + \dot{f}_{\text{em}}^{(n)}(t_1)t, \quad (1)$$

where an overdot denotes a derivative with respect to  $t$  and  $t_1$  labels the time of the first TOA. The stochastic torque is assumed to be a zero-mean, white noise process. Specifically, the frequency evolves according to an Ornstein-Uhlenbeck process, described by a Langevin equation with a time-dependent drift term (Vargas & Melatos 2023)

$$\frac{df_p^{(n)}}{dt} = -\gamma^{(n)}[f_p^{(n)} - f_{\text{em}}^{(n)}(t)] + \dot{f}_{\text{em}}^{(n)}(t) + \xi^{(n)}(t). \quad (2)$$

In Equation (2),  $f_{\text{em}}^{(n)}$  is the solution of the electromagnetic spin-down equation given by Equation (1),  $\dot{f}_{\text{em}}^{(n)}$  is the spin derivative,  $\gamma^{(n)}$  is a damping constant whose reciprocal specifies the mean-reversion timescale, and  $\xi^{(n)}(t)$  is a white noise stochastic process which satisfies:

$$\langle \xi^{(n)}(t) \rangle = 0, \quad (3)$$

$$\langle \xi^{(n)}(t) \xi^{(n')}(t') \rangle = [\sigma^{(n)}]^2 \delta_{n,n'}(t - t'). \quad (4)$$

In Equation (4)  $[\sigma^{(n)}]^2$  is the variance of  $\xi^{(n)}$  and parametrizes the amplitude of the noise. Combined with the mean reversion it gives characteristic root mean square fluctuations  $\approx \sigma^{(n)} / [\gamma^{(n)}]^{1/2}$  in  $f_p^{(n)}(t)$  (Gardiner 2009). It is important to note that white noise fluctuations in  $\xi(t)$  translate into red noise fluctuations in the rotational phase  $\phi(t) = \int_{t_1}^t dt' f_p(t')$  after being filtered by the terms involving  $d/dt$  and  $\gamma$  in Equation (1), consistent with the observed power spectral density of typical millisecond pulsars in the nHz band relevant to PTA experiments.

As discussed in K1 Equations (1)–(4) represent a phenomenological model, which aims to reproduce qualitatively the typical timing behaviour observed in PTAs, viz. a mean-reverting random walk about a secular spin-down trend (Agazie et al. 2023b; Antoniadis et al. 2023b; Zic et al. 2023b). Equations (1)–(4) are not derived from first principles by applying a microphysical theory. As a first pass, they also exclude certain phenomenological elements, which are likely to be present in reality, e.g. the classic, two-component, crust-superfluid structure inferred from post-glitch recoveries (Baym et al. 1969; van Eysden & Melatos 2010; Gügercinoğlu & Alpar 2017; Meyers et al. 2021a,b). An approach akin to Equations (1)–(4) has been followed successfully in other timing analyses in the context of anomalous braking indices (Vargas & Melatos 2023) and hidden Markov model glitch searches (Melatos et al. 2020; Lower et al. 2021; Dunn et al. 2022, 2023). However, Equations (1)–(4) involve significant idealizations, which must be recognized at the outset (Meyers et al. 2021b,a; Vargas & Melatos 2023). First, the white noise driver  $\xi(t)$  in Equation (2) is not differentiable, which makes the formal interpretation of  $d^2 f_p / dt^2$  ambiguous, even though  $d^2 f_p / dt^2$  is not used in the PTA analysis proposed in this paper. Second, the white spectrum assumed for  $\xi(t)$  may or may not be suitable for millisecond pulsars in PTAs. It is challenging observationally to infer the spectrum of  $\xi(t)$  from the observed spectrum of the phase residuals, because the inference is conditional on the (unknown) dynamical model governing  $df_p/dt$ . For small-amplitude fluctuations sampled relatively often, as in millisecond pulsars in PTAs, it is likely that  $\xi(t)$  is white to a good approximation over the inter-TOA intervals and generates red phase residuals as observed, but caution is warranted nevertheless. Third, the Brownian increment  $dB(t) = \xi(t)dt$  does not include non-Gaussian excursions such as Lévy flights (Sornette

2004), which have not been ruled out by pulsar timing experiments to date. The above three idealizations are supplemented by other, physical approximations noted above, e.g. neglecting  $n_{\text{em}}$  in Equation (1) and differential rotation between the crust and superfluid in Equation (2).

## 2.2 Modulation of pulsar frequency by a GW

In the presence of a GW, the pulse frequency measured by an observer in the local rest frame of the neutron star's center of mass is different from that measured by an observer on Earth. Specifically, the pulse frequency is modulated harmonically at the GW frequency. We derive the nonlinear measurement equation relating  $f_m(t)$  to  $f_p(t)$  in this section. The measurement equation is a key input into the Kalman filter in Section (5.1)

### 2.2.1 Plane GW perturbation

We consider a gravitational plane wave from a single, distant source, which perturbs a background Minkowski metric  $\eta_{\mu\nu}$  as

$$g_{\mu\nu} = \eta_{\mu\nu} + H_{\mu\nu} \exp \{i[\Omega(\mathbf{n} \cdot \mathbf{x} - t) + \Phi_0]\}, \quad (5)$$

with spatial coordinates  $\mathbf{x}$  and global coordinate time  $t$ . The GW has a constant angular frequency  $\Omega$ , propagates in the  $\mathbf{n}$ -direction (where  $\mathbf{n}$  is a unit vector), has amplitude tensor  $H_{\mu\nu}$ , and has a phase offset  $\Phi_0$ . Throughout this paper we work with pulsar TOAs which have been defined relative to the Solar System barycentre (SSB). We are free to choose our coordinate system such that  $\Phi_0$  is the GW phase at  $t = 0$  at the SSB. In this paper  $\Omega$  has no time dependence; the source is approximated as monochromatic. Studies of SMBHB inspirals in the PTA context show that the gravitational wave frequency  $f_{\text{gw}}$  ( $= \Omega/2\pi$ ) evolves according to (e.g. [Zhu et al. 2015a](#)),

$$\Delta f_{\text{gw}} \simeq 3.94 \text{ nHz} \left( \frac{M_c}{10^9 M_\odot} \right)^{5/3} \left[ \frac{f_{\text{gw}}(t = t_1)}{10^{-7} \text{ Hz}} \right]^{11/3} \left( \frac{T_{\text{obs}}}{10 \text{ yr}} \right), \quad (6)$$

where  $M_c$  is the chirp mass of the SMBHB,  $f_{\text{gw}}(t = t_1)$  is the GW frequency at the time of the first observation, and  $T_{\text{obs}}$  is the length of time over which  $\Delta f_{\text{gw}}$  is measured, which for PTAs is  $\sim 10$  years. A source can be considered monochromatic, if  $\Delta f_{\text{gw}}$  is less than the PTA frequency resolution  $1/T_{\text{obs}}$ . From Equation (6) we can see that only those binaries which are massive or at high frequency experience significant frequency evolution over typical PTA timespans. The majority of SMBHBs detectable with PTAs are expected to satisfy  $\Delta f_{\text{gw}} < 1/T_{\text{obs}}$ ; for a PTA composed of pulsars with a mean distance of 1.5 kpc, 78% of simulated SMBHBs satisfy this condition for the current IPTA, whilst for the second phase of the Square Kilometer Array this fraction drops to 52%; see Figure 7 in [Rosado et al. \(2015\)](#). We are therefore justified in treating the GW source as monochromatic as a first pass in this introductory paper ([Sesana et al. 2008](#); [Sesana & Vecchio 2010](#); [Ellis et al. 2012](#)).

The amplitude tensor  $H_{\mu\nu}$  has zero temporal components ( $H_{0\mu} = H_{\mu 0} = 0$ ). The spatial part is

$$H_{ij} = h_+ e_{ij}^+(\mathbf{n}) + h_\times e_{ij}^\times(\mathbf{n}), \quad (7)$$

where  $h_+$  and  $h_\times$  are the respective polarisation amplitudes. The plus and cross polarisation tensors  $e_{ij}^+$  and  $e_{ij}^\times$  are uniquely defined by the principal axes of the wave, viz. the unit 3-vectors  $\mathbf{k}$  and  $\mathbf{l}$ , according to

$$e_{ij}^+(\mathbf{n}) = k_i k_j - l_i l_j, \quad (8)$$

$$e_{ij}^\times(\mathbf{n}) = k_i l_j + l_i k_j. \quad (9)$$

The principal axes are in turn specified by the location of the GW source on the sky (colatitude  $\theta$ , azimuth  $\phi$ ) and the polarisation angle  $\psi$  according to

$$\begin{aligned} \mathbf{k} = & (\sin \phi \cos \psi - \sin \psi \cos \phi \cos \theta) \hat{\mathbf{x}} \\ & - (\cos \phi \cos \psi + \sin \psi \sin \phi \cos \theta) \hat{\mathbf{y}} \\ & + (\sin \psi \sin \theta) \hat{\mathbf{z}}, \end{aligned} \quad (10)$$

$$\begin{aligned} \mathbf{l} = & (-\sin \phi \sin \psi - \cos \psi \cos \phi \cos \theta) \hat{\mathbf{x}} \\ & + (\cos \phi \sin \psi - \cos \psi \sin \phi \cos \theta) \hat{\mathbf{y}} \\ & + (\cos \psi \sin \theta) \hat{\mathbf{z}}, \end{aligned} \quad (11)$$

where e.g.  $\hat{\mathbf{x}}$  is a unit vector in the direction of the  $x$ -axis. The direction of GW propagation is related to the principal axes by

$$\mathbf{n} = \mathbf{k} \times \mathbf{l}. \quad (12)$$

### 2.2.2 Measurement equation

In general radio pulses from a pulsar are transmitted as amplitude modulations of a radio-frequency carrier wave. They are described by the geometric object  $\mathbf{p}$  which we identify as the 4-momentum of the radio pulse. The presence of a GW induces a shift in the temporal component of the covariant 4-momentum, while the photon travels from the emitter to the observer, i.e.  $\Delta p_t = p_t|_{\text{observer}} - p_t|_{\text{emitter}}$ . One obtains (e.g. [Maggiore 2018](#))

$$\Delta p_t = \frac{\pi f_p h_{ij}(t; \mathbf{x} = 0) q^i q^j}{1 + \mathbf{n} \cdot \mathbf{q}} \left[ 1 - e^{i\Omega(1 + \mathbf{n} \cdot \mathbf{q})d} \right], \quad (13)$$

where  $f_p$  is the pulse frequency measured in the momentarily comoving reference frame of an observer. We write  $h_{ij} = g_{ij} - \eta_{ij}$ ,  $\mathbf{q}$  is the unit vector connecting the observer and the pulsar and  $d$  is the distance to the pulsar. We take the pulsar location to be constant i.e. neither  $\mathbf{q}$  nor  $d$  are functions of time. In practice the pulsar locations vary with respect to the Earth but are constant with respect to the SSB. The barycentering correction is typically applied when generating TOAs, e.g. with TEMPO2 ([Hobbs et al. 2006](#)) and related timing software, and is inherited by the frequency time series. Some pulsars, including some PTA pulsars, do have non-negligible proper motions  $O(100) \text{ km s}^{-1}$  after the barycentering corrections have been applied (e.g. [Jankowski et al. 2018](#)), but we do not consider this effect in this paper. Generally the measured frequency of a photon recorded by an observer who is travelling with 4-velocity  $u$  is given by the coordinate-independent expression  $p_\alpha u^\alpha$ . After barycentering, one has  $u^\alpha = (1, 0, 0, 0)$  for both the emitter and the observer to leading order in the respective momentarily comoving reference frames,

$$u^\alpha|_{\text{emitter}} = u^\alpha|_{\text{observer}} = (1, 0, 0, 0), \quad (14)$$

Hence Equation (13) can be written for the  $n$ -th pulsar as

$$f_m^{(n)}(t) = f_p^{(n)}(t - d^{(n)}) g^{(n)}(t) + \varepsilon^{(n)}. \quad (15)$$

where  $d^{(n)}$  labels the distance to the  $n$ -th pulsar, and  $\varepsilon^{(n)}$  is a Gaussian measurement noise which satisfies:

$$\langle \varepsilon^{(n)}(t) \rangle = 0, \quad (16)$$

$$\langle \varepsilon^{(n)}(t) \varepsilon^{(n')}(t') \rangle = \sigma_m^2 \delta_{n,n'}(t - t'). \quad (17)$$

where  $\sigma_m$  is the covariance of the measurement noise and is shared between all pulsars. The measurement function  $g^{(n)}(t)$  is

$$g^{(n)}(t) = 1 - \frac{h_{ij}(t; \mathbf{x} = 0) [q^{(n)}]^i [q^{(n)}]^j}{2[1 + \mathbf{n} \cdot \mathbf{q}^{(n)}]} \left\{ 1 - e^{i\Omega[1 + \mathbf{n} \cdot \mathbf{q}^{(n)}]d^{(n)}} \right\}, \quad (18)$$

and  $[q^{(n)}]^i$  labels the  $i$ -th coordinate component of the  $n$ -th pulsar's position vector  $\mathbf{q}^{(n)}$ . It is also instructive to express Equation (18) in a trigonometric form as,

$$g^{(n)}(t) = 1 - \frac{H_{ij}[q^{(n)}]^i[q^{(n)}]^j}{2[1 + \mathbf{n} \cdot \mathbf{q}^{(n)}]} \times \left[ \cos(-\Omega t + \Phi_0) - \cos\left\{-\Omega t + \Phi_0 + \Omega \left[1 + \mathbf{n} \cdot \mathbf{q}^{(n)}\right] d^{(n)}\right\} \right]. \quad (19)$$

Equations (15)–(19) define a non-linear measurement equation that relates the intrinsic pulsar spin frequency to the pulse frequency measured by an observer on Earth.

### 3 INCLUSION OF THE PULSAR TERM

A naive approach is challenging - look at this noisy likelihood curve  
A reparameterisation is good

### 4 REPRESENTATIVE EXAMPLE: SINGLE SMBHB SOURCE AND COMPARISON WITH EARTH TERMS

### 5 SIGNAL TRACKING, PARAMETER ESTIMATION AND MODEL SELECTION

The set of static parameters  $\theta$  of the model outlined in Section 2 can be separated into parameters controlling the intrinsic frequency evolution of the pulsars in the array and the GW source, i.e.

$$\theta = \theta_{\text{psr}} \cup \theta_{\text{gw}}, \quad (20)$$

with

$$\theta_{\text{psr}} = \left\{ \gamma^{(n)}, \sigma^{(n)}, f_{\text{em}}^{(n)}(t_1), \dot{f}_{\text{em}}^{(n)}(t_1), d^{(n)} \right\}_{1 \leq n \leq N}, \quad (21)$$

and

$$\theta_{\text{gw}} = \{h_0, \iota, \delta, \alpha, \psi, \Omega, \Phi_0\}, \quad (22)$$

where  $\delta$ ,  $\alpha$ , and  $\iota$  are the declination, right ascension and inclination of the GW source respectively<sup>1</sup>. In Equation (22) we reparameterize the two GW polarisation amplitudes,  $h_+$  and  $h_\times$ , in terms of  $\iota$  and the characteristic wave strain  $h_0$  through

$$h_+ = h_0(1 + \cos^2 \iota), \quad (23)$$

$$h_\times = -2h_0 \cos \iota. \quad (24)$$

We use the parametrisation in terms of  $h_0$  and  $\iota$  throughout the remainder of this work. A PTA containing  $N$  pulsars comprises  $7 + 5N$  parameters to estimate. Typically the pulsar parameters are constrained better *a priori* by electromagnetic observations than the GW parameters. For example estimates of pulsar distances are accurate to  $\sim 10\%$  (Cordes & Lazio 2002; Verbiest et al. 2012; Desvignes et al. 2016; Yao et al. 2017), but we have no prior information about  $\delta$  and  $\alpha$ .

In K1

The workflow presented in K1, which we also follow here proceeds as follows

In this section we present a new method to infer  $\theta$  and calculate the marginal likelihood (i.e. the model evidence). In Section 5.1 we

outline how noisy measurements of the pulsar frequency,  $f_m^{(n)}(t)$ , can be used to estimate the hidden state sequence,  $f_p^{(n)}(t)$ , using a Kalman filter. In Section 5.2 we demonstrate how to deploy the Kalman filter in conjunction with a nested sampling technique to perform Bayesian inference. Model selection and the specification of the null model are described in Section 5.3. A complete summary of the workflow is presented in Section 5.4. The method complements traditional PTA analyses because it does not assume a specific functional form (e.g. power law) for the power spectral density of the TOA fluctuations.

### 5.1 Kalman filter and likelihood

The Kalman filter (Kalman 1960) is a Gauss-Markov model used to algorithmically recover a temporal sequence of stochastically evolving system state variables,  $\mathbf{X}(t)$ , which are not observed directly, given a temporal sequence of noisy measurements,  $\mathbf{Y}(t)$ . It finds common use in engineering applications and has been applied successfully in neutron star astrophysics (e.g. Meyers et al. 2021a,b; Melatos et al. 2023). In this work we use the linear Kalman filter, which assumes a linear relation between  $d\mathbf{X}/dt$  and  $\mathbf{X}(t)$  (dynamics) and between  $\mathbf{Y}(t)$  and  $\mathbf{X}(t)$  (measurement), with  $\mathbf{X}(t) = \{f_p^{(n)}(t)\}$  and  $\mathbf{Y}(t) = \{f_m^{(n)}(t)\}$ . Extension to non-linear problems is straightforward using either an extended Kalman filter (Zarchan & Musoff 2000), unscented Kalman filter (Wan & Van Der Merwe 2000) or particle filter (Simon 2006). Equations (15) and (18) are non-linear in the static parameters (e.g.  $d^{(n)}$ ), even though they are linear in  $\mathbf{X}(t)$  and  $\mathbf{Y}(t)$ . Hence inferring the static parameters is a non-linear exercise, to be tackled separately after the linear Kalman filter operates on the data for a fixed set of static parameters. Inference of the static parameters in Equations (15) and (18) by nested sampling is discussed in Section 5.2.

Implementation of the linear Kalman filter for the PTA state-space model in Section 3.2, including the full set of Kalman recursion relations is presented in Appendix A. At each discrete timestep indexed by  $1 \leq i \leq M$ , the Kalman filter returns an estimate of the state variables,  $\hat{\mathbf{X}}_i = \hat{\mathbf{X}}(t_i)$ , and the covariance of those estimates,  $\mathbf{P}_i = \langle \hat{\mathbf{X}}_i \hat{\mathbf{X}}_i^T \rangle$ , where the superscript T denotes the matrix transpose. The filter tracks the error in its predictions of  $\mathbf{X}_i$  by converting  $\hat{\mathbf{X}}_i$  into predicted measurements  $\hat{\mathbf{Y}}_i$  via Equations (15) and (18) and comparing with the actual, noisy measurements  $\mathbf{Y}_i$ . This defines a residual  $\epsilon_i = \mathbf{Y}_i - \hat{\mathbf{Y}}_i$ , which is sometimes termed the innovation. The Kalman filter also calculates the uncertainty in  $\epsilon_i$  via the innovation covariance  $\mathbf{S}_i = \langle \epsilon_i \epsilon_i^T \rangle$ . The innovation and the innovation covariance are then used to correct the state estimate  $\hat{\mathbf{X}}_i$  according to Equation (A8). For a fixed set of static parameters, the Kalman filter returns an estimate of the the state sequence  $\hat{\mathbf{X}}_1, \dots, \hat{\mathbf{X}}_M$  which minimizes the mean square error. We explain how to use this intermediate output to infer the optimal values of the static parameters  $\theta$  in Section 5.2

The Gaussian log-likelihood of obtaining  $\mathbf{Y}_i$  given  $\hat{\mathbf{X}}_i$  can then be calculated at each timestep from the Kalman filter output according to

$$\log \mathcal{L}_i = -\frac{1}{2} \left( N \log 2\pi + \log |\mathbf{S}_i| + \epsilon_i^T \mathbf{S}_i^{-1} \epsilon_i \right). \quad (25)$$

The total log-likelihood for the entire sequence is

$$\log \mathcal{L} = \sum_{i=1}^M \log \mathcal{L}_i. \quad (26)$$

<sup>1</sup>  $\iota$  is the angle between the normal to the SMBHB orbital plane,  $\mathbf{L}$ , and the observer's line of sight, i.e.  $\cos \iota = \mathbf{n} \cdot \mathbf{L}$ .



Given  $\mathbf{Y}_1, \dots, \mathbf{Y}_M$ ,  $\mathcal{L}$  is a function of the estimates  $\hat{\theta}$  of the static parameters passed to the Kalman filter, i.e.  $\mathcal{L} = \mathcal{L}(\mathbf{Y}|\hat{\theta})$ . Similarly the estimates of the state and measurement variables,  $\hat{\mathbf{X}}$  and  $\hat{\mathbf{Y}}$ , are functions of  $\hat{\theta}$ . If  $\hat{\theta}$  is close to the true underlying parameters  $\theta$ , then the errors in  $\hat{\mathbf{X}}$  and  $\hat{\mathbf{Y}}$  are minimized and  $\mathcal{L}$  is maximised. This is illustrated with synthetic data in Figure ?? . In the left column, a time series of  $f_m^{(1)}(t)$  including Gaussian noise (middle panel, red curve) is generated from Equations (1)–(4), (16), and (17) for a single pulsar and fed into the Kalman filter along with the true static parameters  $\hat{\theta} = \theta$ . The Kalman filter recovers the evolution of  $f_p^{(1)}(t)$  with high fidelity; the estimate of  $f_p^{(1)}(t)$  (left top panel, blue curve) overlaps almost perfectly with the true  $f_p^{(1)}(t)$  (left top panel, green curve). The predicted state  $\hat{f}_p^{(1)}(t)$  is converted into a predicted measurement  $\hat{f}_m^{(1)}(t)$  (middle panel, magenta curve), which again overlaps almost perfectly with the true measurement. The residuals  $\epsilon(t)$  between the true and predicted measurements are small ( $\lesssim 0.1\%$ ) and normally distributed (left bottom panel). By contrast, in the right column, the exercise is repeated while passing slightly incorrect static parameters to the Kalman filter, where  $\Omega$  has been perturbed from its true value by 20%. In this case the Kalman filter fails to track  $f_p^{(1)}(t)$  accurately, as the discrepancy between the blue and green curves in the top panel of the right-hand column indicates. It similarly fails to predict  $f_m^{(1)}(t)$  accurately, as shown by the discrepancy between the red and magenta curves in the middle panel, and the residuals are no longer distributed normally (right bottom panel).

## 5.2 Nested sampling

We can use the likelihood returned by the Kalman filter, Equation (26), in conjunction with likelihood-based inference methods to estimate the posterior distribution of  $\theta$  by Bayes' Rule,

$$p(\theta|\mathbf{Y}) = \frac{\mathcal{L}(\mathbf{Y}|\theta) \cdot \pi(\theta)}{\mathcal{Z}}, \quad (27)$$

where  $\pi(\theta)$  is the prior distribution on  $\theta$  and  $\mathcal{Z}$  is the marginalised likelihood, or evidence

$$\mathcal{Z} = \int d\theta \mathcal{L}(\mathbf{Y}|\theta) \pi(\theta). \quad (28)$$

We estimate the posterior distribution and the model evidence through nested sampling (Skilling 2006) in this paper. Nested sampling evaluates marginalised likelihood integrals, of the form given by Equation (28), which also approximates the posterior by returning samples from  $p(\theta|\mathbf{Y})$ . It does so by drawing a set of  $n_{\text{live}}$  live points from  $\pi(\theta)$  and iteratively replacing the live point with the lowest likelihood with a new live point drawn from  $\pi(\theta)$ , where the new live point is required to have a higher likelihood than the discarded point. The primary advantage of nested sampling is its ability to compute  $\mathcal{Z}$ , on which model selection relies. Nested sampling is also computationally efficient and can handle multi-modal problems (Ashton et al. 2022). For these reasons, it has enjoyed widespread adoption in the physical sciences, particularly within the cosmological community (Mukherjee et al. 2006; Feroz & Hobson 2008; Handley et al. 2015), neutron star astrophysics (Meyers et al. 2021a,b; Melatos et al. 2023), particle physics (Trassinelli 2019) and materials science (Pártay et al. 2009). For a review of nested sampling we refer the reader to Buchner (2021a) and Ashton et al. (2022). Multiple nested sampling algorithms and computational libraries exist. (e.g. Feroz & Hobson 2008; Feroz et al. 2009; Handley et al. 2015; Speagle 2020; Buchner 2021b).

In gravitational wave research it is common to use the dynesty sampler (Speagle 2020) via the Bilby (Ashton & Talbot 2021) front-end library. We follow this precedent and use Bilby for all nested sampling Bayesian inference in this work.

The primary tunable parameter in nested sampling is  $n_{\text{live}}$ . A greater number of live points is advantageous for large parameter spaces and multi-modal problems, whilst the uncertainties in the evidence and the posterior scale as  $O(n_{\text{live}}^{-1/2})$ . However the computational runtime scales as  $O(n_{\text{live}})$ . Ashton et al. (2022) offered a rule-of-thumb trade-off, where the minimum number of live points should be greater than the number of static parameters. Informal empirical tests conducted as part of this paper support the trade-off suggested by Ashton et al. (2022); we find typically that the true  $\theta$  is contained within the 90% credible interval of the derived 1D marginalised posteriors of  $\hat{\theta}$  for  $n_{\text{live}} > 7 + 5N$  with  $N \leq 50$ . Unless stated otherwise we take  $n_{\text{live}} = 1000$  for all results presented in this work.

## 5.3 Model selection

The evidence integral  $\mathcal{Z}$  returned by nested sampling is the probability of the data  $\mathbf{Y}$  given a model  $\mathcal{M}_i$ . We compare competing models via a Bayes factor,

$$\beta = \frac{\mathcal{Z}(\mathbf{Y}|\mathcal{M}_1)}{\mathcal{Z}(\mathbf{Y}|\mathcal{M}_0)}. \quad (29)$$

Throughout this work we take  $\mathcal{M}_1$  to be the state-space model that includes the presence of a GW.  $\mathcal{M}_0$  is the null model, which assumes no GW exists in the data. This is equivalent to setting  $g^{(n)}(t) = 1$  in Equations (18) and (19). The Bayes factors we quote in this work therefore quantify whether the data supports evidence for a GW compared to there being no GW signal present.

## 5.4 Summary of workflow

For the reader's convenience we now summarise the workflow for a representative PTA analysis using the Kalman filter and nested sampler for parameter estimation and model selection:

- (i) Specify a PTA composed of  $N$  pulsars
- (ii) Obtain  $N$  data inputs  $f_m^{(n)}(t)$ , collectively labelled  $\mathbf{Y}$
- (iii) Specify a state-space model  $\mathcal{M}$ , with static parameters  $\theta$ .
- (iv) Specify prior distribution  $\pi(\theta)$
- (v) Sample  $n_{\text{live}}$  points from  $\pi(\theta)$
- (vi) For each live point:
  - (a) Pass the sample  $\theta_{\text{sample}}$  to the Kalman filter.
  - (b) Iterate over the input data using the Kalman filter and obtain a single  $\log \mathcal{L}$  value, Equation (26)
- (vii) Remove the live point with the lowest likelihood value,  $\log \mathcal{L}_{\text{lowest}}$
- (viii) Constrained sample a new live point from  $\pi(\theta)$ , subject to the requirement that the new likelihood  $\log \mathcal{L}_{\text{new}} > \log \mathcal{L}_{\text{lowest}}$ , where  $\log \mathcal{L}_{\text{new}}$  is calculated via steps (a)–(b).
- (ix) Nested sampler updates  $p(\theta|\mathbf{Y})$  and  $\mathcal{Z}$
- (x) Repeat steps (vii) – (ix) until convergence criteria are satisfied.

In order to compute  $\beta$  the above workflow is repeated for a different  $\mathcal{M}$ . The resulting  $\mathcal{Z}$  values can then be compared. We remind the reader that the above workflow differs from a realistic PTA analysis

in an important respect, namely that the data are input as frequency time series  $f_m^{(n)}(t)$  instead of pulse TOAs. The generalization to TOAs is subtle and will be tackled in a forthcoming paper.

## 6 VALIDATION PROCEDURE WITH SYNTHETIC DATA

In this section we outline how synthetic data is generated in order to validate the analysis scheme in Section 5. In Section 6.1 we describe how to construct a representative synthetic PTA, and how to set astronomically reasonable values for the static pulsar parameters  $\theta_{\text{psr}}$ . In Section 6.2 we demonstrate how to solve the equations of Section 2 for the synthetic PTA so as to generate noisy, frequency time series  $f_m^{(n)}(t)$ .

### 6.1 Constructing a synthetic PTA

We consider, by way of illustration, a synthetic PTA composed of the 47 pulsars in the 12.5 year NANOGrav dataset (Arzoumanian et al. 2020b). The NANOGrav pulsars are chosen arbitrarily as being representative of a typical PTA; the analysis below extends unchanged to any other PTA. For each pulsar we adopt the fiducial values for  $q^{(n)}$ ,  $d^{(n)}$ ,  $f_{\text{em}}^{(n)}(t_1)$ , and  $\dot{f}_{\text{em}}(t_1)$ , with the latter two quantities evaluated at the Solar System barycenter. A table of fiducial values is presented in Appendix B for the sake of reproducibility. The sky positions and colour-coded distances of the pulsars are displayed in Figure ???. The pulsar parameters are acquired via the Australia Telescope National Facility (ATNF) pulsar catalogue (Manchester et al. 2005) using the `psrqpy` package (Pitkin 2018). The remaining static pulsar parameters are  $\gamma^{(n)}$  and  $\sigma^{(n)}$ , for which no direct measurements exist. The ratio  $\sigma^{(n)}/[\gamma^{(n)}]^{1/2}$  sets the typical root mean square fluctuations in  $f_p^{(n)}(t)$ , as discussed in Section 2.1 and the mean reversion timescale typically satisfies  $[\gamma^{(n)}]^{-1} \gg T_{\text{obs}}$  (Price et al. 2012; Meyers et al. 2021a,b; Vargas & Melatos 2023). In this paper, for the sake of simplicity in the absence of independent measurements, we fix  $\gamma^{(n)} = 10^{-13} \text{ s}^{-1}$  for all  $n$ . We follow two complementary approaches in order to set physically reasonable values for  $\sigma^{(n)}$ . The first approach relies on the empirical timing noise model from Shannon & Cordes (2010b) which gives the standard deviation of the pulsar TOAs,  $\sigma_{\text{TOA}}^{(n)}$ , empirically as

$$\ln\left(\frac{\sigma_{\text{TOA}}^{(n)}}{\mu\text{s}}\right) = \ln\alpha_1 + \alpha_2 \ln f_p^{(n)} + \alpha_3 \ln\left(\frac{\dot{f}_p^{(n)}}{10^{-15}\text{s}^{-2}}\right) + \alpha_4 \ln\left(\frac{T_{\text{cad}}}{1\text{ year}}\right), \quad (30)$$

where  $T_{\text{cad}}$  is the cadence of the timing observations and for MSPs the best fit parameters are  $\ln\alpha_1 = -20 \pm 20$ ,  $\alpha_2 = 1 \pm 2$ ,  $\alpha_3 = 2 \pm 1$ ,  $\alpha_4 = 2.4 \pm 0.6$ , for  $\pm 2\sigma$  confidence limits. The uncertainties on these parameters are broad; for the purposes of generating astrophysically representative synthetic data we straightforwardly take the best fit parameter value. Throughout this work we assume that all pulsars are observed with a weekly cadence,  $T_{\text{cad}} = 1$  week. In order to relate Equation (30), to  $\sigma^{(n)}$  in Equation (1), we equate  $\sigma^{(n)}$  to the root mean square TOA noise accumulated over one week, obtaining

$$\sigma^{(n)} \approx \sigma_{\text{TOA}}^{(n)} f_p^{(n)} T_{\text{cad}}^{-3/2}. \quad (31)$$

For the synthetic NANOGrav PTA depicted in Figure ??, the median  $\sigma^{(n)}$  calculated in this way is  $\sigma^{(n)} = 5.51 \times 10^{-24} \text{ s}^{-3/2}$ ,  $\min[\sigma^{(n)}] = 1.67 \times 10^{-26} \text{ s}^{-3/2}$  for PSR J0645+5158 and

$$\max[\sigma^{(n)}] = 2.56 \times 10^{-19} \text{ s}^{-3/2} \text{ for PSR J1939+2134}$$

As a cross-check, we estimate  $\sigma^{(n)}$  by solving Equation (1) numerically using the `baboo` package<sup>2</sup>, generating a synthetic phase solution,

$$\varphi^{(n)}(t) = \int_0^t dt' f_p^{(n)}(t'), \quad (32)$$

and adjusting  $\sigma^{(n)}$  iteratively to generate phase residuals which qualitatively (i.e. visually) resemble empirical phase residuals measured from real pulsars; see Section 2 in Vargas & Melatos (2023) for a successful example of this approach. We obtain empirical phase residuals from the NANOGrav 12.5 year wideband timing dataset (Pennucci & Collaboration 2020; Alam et al. 2021). A visual cross-check is sufficient for the purposes of this paper, where we seek broadly representative values for  $\sigma^{(n)}$ . In an analysis involving real PTA data,  $\sigma^{(n)}$  would be estimated from the data jointly with the other static parameters in  $\theta$ . In Figure ?? we compare the synthetic and empirical residuals for PSR J1939+2134, one of the 47 NANOGrav pulsars plotted in Figure ??. We see that the synthetic and empirical residuals are qualitatively similar. For this pulsar the synthetic residuals were generated using  $\sigma^{(n)} = 10^{-27} \text{ s}^{-3/2}$  which is  $\ll$  than the value inferred from the noise model of Equations (30) and (31), using the best fit parameters,  $\sigma^{(n)} = 2.56 \times 10^{-19} \text{ s}^{-3/2}$ . This highlights how given the large uncertainties on the parameters of Equation (30) using the best fit values may be less appropriate for certain pulsars. If we set  $\ln\alpha_1 = -38$  in Equation (30), close to its lower limit, then the estimated  $\sigma^{(n)}$  for PSR J1939+2134 is  $\sigma^{(n)} = 3.90 \times 10^{-27}$  which is in much better agreement with the value inferred from the residuals comparison method. This discrepancy between the two estimates of  $\sigma^{(n)}$  is particular to PSR J1939+2134; for all other pulsars we find the two estimates of  $\sigma^{(n)}$  are comparable to within a few orders of magnitude.

### 6.2 Generating a synthetic sequence of pulse frequencies

We generate  $N$  synthetic noisy time series of the measured pulse frequency  $f_m^{(n)}(t)$ , one for each pulsar  $1 \leq n \leq N$ , as follows:

- (i) Integrate Equations (1)–(4) numerically for the synthetic PTA described in Section 6.1, to obtain random realizations of  $f_p^{(n)}(t)$  for  $1 \leq n \leq N$ .
- (ii) Map from  $f_p^{(n)}(t)$  to  $f_m^{(n)}(t)$  via Equations (15) and (18)

Equations (1)–(4) are solved by a Runge-Kutta Itô integrator implemented in the `sdeint` python package<sup>3</sup>. The static pulsar parameters  $\theta_{\text{psr}}$  are completely specified for the synthetic PTA outlined in Section (6.1). For this work we consider all pulsars to be observed for  $T_{\text{obs}} = 10$  years, uniformly sampled with a weekly cadence. The measurement noise covariance as defined in Equation (17) can be approximately related to the uncertainty in the pulse TOA,  $\sigma_{\text{TOA}}$ , as

$$\sigma_m \approx f_p^{(n)} \sigma_{\text{TOA}} T_{\text{cad}}^{-1}. \quad (33)$$

Although Equations (31) and (33) superficially resemble one another, they are distinct; Equation (31) deals with the timing noise intrinsic to the pulsar due to rotational irregularities, whereas Equation (33) handles the detector measurement noise. For a millisecond pulsar with  $f_p^{(n)} \sim 100 \text{ Hz}$  observed with a weekly

<sup>2</sup> <https://github.com/meyers-academic/baboo>

<sup>3</sup> <https://github.com/mattja/sdeint>

cadence and  $\sigma_{\text{TOA}} \sim 1\mu\text{s}$  Equation (33) implies  $\sigma_m \sim 10^{-10}$  Hz. The most accurately timed pulsars might have  $\sigma_{\text{TOA}} \sim 10$  ns or  $\sigma_m \sim 10^{-12}$  Hz. Throughout this paper we fix  $\sigma_m = 10^{-11}$  Hz and take it as known *a priori* rather than a parameter to be inferred. Note that whilst  $\sigma_m$  is the same for every pulsar,  $f_m^{(n)}$  is modified by a different realisation of the measurement noise for each pulsar.

Finite precision arithmetic leads to numerical errors when solving Equations (1)–(4) in the regime  $\sigma^{(n)} \Delta W_t \ll f_p^{(n)}$  relevant to PTAs, where  $\Delta W_t$  labels an increment of Brownian motion (see Appendix A). To fix the problem, we subtract the deterministic frequency evolution and track the new variable, equivalent to a change of variables:

$$f_p^{*(n)} = f_p^{(n)} - f_{\text{em}}^{(n)}. \quad (34)$$

We similarly modify the measurement variable to be

$$f_m^{*(n)} = f_m^{(n)} - f_{\text{em}}^{\diamond(n)}, \quad (35)$$

where  $f_{\text{em}}^{\diamond(n)}$  is a guess of the deterministic spin-down based on the pulsar ephemeris which is measured to high accuracy in practice by TEMPO2. For synthetic data we can set  $f_{\text{em}}^{\diamond(n)} = f_{\text{em}}^{(n)}$  without loss of generality, but this is impossible generally for astronomical observations, because the spin-down ephemeris is only known approximately. Equation (15) is then updated to read

$$f_m^{*(n)}(t) = f_p^{*(n)}(t-d)g^{(n)}(t) - f_{\text{em}}^{(n)}(t-d) \left[1 - g^{(n)}(t)\right]. \quad (36)$$

We emphasise that this change of variables is a convenient device to bring the numerical values into a reasonable dynamic range without having to use excessively long floating point formats (e.g. long double, quadruple). It does not remove any degrees of freedom nor does it involve an approximation. In particular  $f_{\text{em}}^{(n)}(t_1)$  and  $\dot{f}_{\text{em}}^{(n)}(t_1)$  remain as static parameters but appear in the measurement equation (36) rather than the dynamical state equations, Equations 1–(4).

## 7 REPRESENTATIVE EXAMPLE: SINGLE SMBHB SOURCE

In this Section we apply the analysis scheme of Section 5 and the validation procedure of Section 6 to a PTA which is perturbed by a GW from an individual quasi-monochromatic SMBHB source. We remind the reader that the analysis in this paper targets a single SMBHB source. The analysis of a stochastic background composed of the superposition of multiple sources is more challenging and is postponed to a forthcoming paper. In Section 7.1 we outline the priors chosen on  $\theta$ . In Section 7.2 we apply the workflow of Section 5.4 to estimate  $\theta$ , for a single realisation of the pulsar process noise and the measurement noise. We go on in Section ?? to extend this parameter estimation procedure to multiple realisations of the noise processes. In Section 7.3 we calculate the the detectability of the source as a function of  $h_0$ .

The static GW source parameters  $\theta_{\text{gw}}$  used for this injection are selected to be astrophysically reasonable and representative. The injected  $\theta_{\text{gw}}$  are summarised in the “Injected Values” column of Table 1. The static pulsar parameters  $\theta_{\text{psr}}$  are as described in Section 6.1 and are also shown in Table 1.

### 7.1 Prior distributions

For the Bayesian methods used in this work, one must select reasonable priors,  $\pi(\theta)$ , for the complete set complete of static parameters. For  $\pi(\theta_{\text{gw}})$  we choose standard non-informative priors (e.g. Bhagwat et al. 2021) as summarised in Table 1. The choice of  $\pi(\theta_{\text{psr}})$  requires some additional discussion.

The parameters which govern the deterministic evolution of the pulsar spin frequency,  $f_{\text{em}}^{(n)}(t_1)$  and  $\dot{f}_{\text{em}}^{(n)}(t_1)$  are well-determined by existing radio timing observations. We can identify  $f_{\text{em}}^{(n)}(t_1)$  with the pulsar barycentric rotation frequency and  $\dot{f}_{\text{em}}^{(n)}(t_1)$  with the time derivative of barycentric rotation frequency quoted by pulsar observation catalogues. For the 12.5 year NANOGrav pulsars the median error in the barycentric rotation frequency is  $\sim 7 \times 10^{-13}$  Hz, and in the time derivative of barycentric rotation frequency is  $\sim 1.8 \times 10^{-20} \text{ s}^{-2}$ . Whilst it is then clear that  $f_{\text{em}}^{(n)}(t_1)$ ,  $\dot{f}_{\text{em}}^{(n)}(t_1)$  are typically measured with very high precisions, for this work we adopt a broader uniform prior,  $\pm 10^3 \epsilon_a^{(n)}$  either side of the true value, where  $\epsilon_a^{(n)}$  is the error for that pulsar quoted in the ATNF catalogue and the subscript  $a = \{f, \dot{f}\}$  identifies the error as being in  $f_{\text{em}}^{(n)}(t_1)$ , or  $\dot{f}_{\text{em}}^{(n)}(t_1)$  respectively. In this way we can test how well our method performs without requiring exceptionally precise and accurate measurements of the pulsar parameters to be made *a priori*. Instead we consider the pulsar parameters similarly to the GW source parameters and estimate them optimally within a consistent framework. Adopting looser priors in this way allows us to subject our method to a more stringent test and confirm that we still accurately estimate GW parameters without highly accurate and precise initial estimates of  $f_{\text{em}}^{(n)}(t_1)$  and  $\dot{f}_{\text{em}}^{(n)}(t_1)$ .

The pulsar distances  $d^{(n)}$  are typically not as well constrained as  $f_{\text{em}}^{(n)}(t_1)$  and  $\dot{f}_{\text{em}}^{(n)}(t_1)$ , with uncertainties typically on the order of  $\sim 10\%$  (Arzoumanian et al. 2018a; Yao et al. 2017). However for this work, whilst the pulsar distance is used when generating synthetic data, as is discussed in Section 7.2 the pulsar distance is not used for our inference model and so we do not need to set a prior on  $d^{(n)}$ . Similarly we do not set a prior on  $\gamma^{(n)}$ ; since  $[\gamma^{(n)}]^{-1} \gg T_{\text{obs}}$  this means that  $\gamma^{(n)}$  is effectively “unobservable” over the decadal timescales that we are interested in. That is, for  $T_{\text{obs}} = 10$  years the solution of Equation (1) is effectively independent of the choice of  $\gamma^{(n)}$  as long as the condition  $[\gamma^{(n)}]^{-1} \gg T_{\text{obs}}$  is satisfied. It is therefore sufficient to consider  $\gamma^{(n)}$  to be known *a priori* and set it at its true injected value. We briefly explored setting an uninformative prior on  $\gamma^{(n)}$  over e.g. LogUniform( $10^{-15}$ ,  $10^{-10}$ )  $\text{s}^{-1}$  as well as setting  $\gamma^{(n)}$  at some fixed value away from the true injected value (e.g. set  $\gamma^{(n)} = 10^{-14} \text{s}^{-1}$  rather than  $10^{-13} \text{s}^{-1}$ ) but the results are unchanged. Regarding  $\sigma^{(n)}$ , the majority of pulsars in our synthetic PTA have  $\sigma^{(n)} \sim 10^{-20} - 10^{-22} \text{s}^{-3/2}$  as calculated from Equations (30) and (31). For these pulsars we set an uninformative broad prior of LogUniform( $10^{-23}$ ,  $10^{-19} \text{s}^{-3/2}$ ). The single exception is PSR J1939+2134 which has a particularly large  $\dot{f}_{\text{em}}^{(n)}(t_1)$  compared to the other pulsars in the array and so  $\sigma \sim 10^{-16} \text{s}^{-3/2}$ . For our artificial dataset we set  $\sigma = 10^{-21} \text{s}^{-3/2}$  for this pulsar in order to bring the process noise for all pulsars into a consistent range that can be described by a single constrained prior.

Since we are not setting priors on  $\gamma^{(n)}$  or  $d^{(n)}$  this reduces the dimensionality of the parameter space to  $7 + 3N$ . We use the notation  $\theta_{\text{psr, reduced}}$  to refer to the reduced parameter space, c.f. Equation 21.

Explicitly,

$$\theta_{\text{psr, reduced}} = \left\{ f_{\text{em}}^{(n)}(t_1), \dot{f}_{\text{em}}^{(n)}(t_1), \sigma^{(n)} \right\}_{1 \leq n \leq N}. \quad (37)$$

All the injected static parameters and their corresponding priors for this representative example are summarised in Table 1.

## 7.2 Parameter estimation

In this section we derive the posterior probability distributions for the static parameters. However, it is first necessary to discuss a modification to the general measurement equation used in the inference model, Equation (19). It can be seen that the equation generally separates into two cosine terms. The first term,  $\cos(-\Omega t + \Phi_0)$ , depends only on the GW source parameters and is shared across all pulsars. The argument of the cosine corresponds to the GW phase at the observer on Earth. Conversely the second term,  $\cos\left(-\Omega t + \Phi_0 + \Omega \left(1 + \mathbf{n} \cdot \mathbf{q}^{(n)}\right) d\right)$ , also depends on  $d^{(n)}$  and  $\mathbf{q}^{(n)}$  and so varies between pulsars. The argument of the cosine corresponds to the GW phase at each the individual pulsar. The first and second terms are commonly referred to as the “Earth term” and “pulsar term” respectively. Whilst the Earth term is phase coherent between all pulsars and so can be summed across the array to increase the total signal-to-noise, the pulsar terms each have individual phases and in standard PTA analysis are typically considered as a source of self noise and dropped from the analysis (e.g. Sesana & Vecchio 2010; Babak & Sesana 2012a; Petiteau et al. 2013; Zhu et al. 2015b; Taylor et al. 2016; Goldstein et al. 2018; Charisi et al. 2023) at the expense of a modestly reduced detection probability ( $\sim 5\%$ ) and the introduction of a bias in the sky localisation (Zhu et al. 2016; Chen & Wang 2022). In this work we follow the standard analysis approach and drop the pulsar terms from our model. Explicitly the measurement equation used in the Kalman filter is

$$f_{\text{m}}^{(n)}(t) = f_{\text{p}}^{(n)}(t - d) g_{\text{Earth}}^{(n)}(t), \quad (38)$$

with

$$g_{\text{Earth}}^{(n)}(t) = 1 - \frac{H_{ij} [q^{(n)}]^i [q^{(n)}]^j}{2[1 + \mathbf{n} \cdot \mathbf{q}^{(n)}]} \cos(-\Omega t + \Phi_0). \quad (39)$$

We defer the inclusion of the pulsar terms to a future work. As discussed in Section 7.1 this choice reduces the dimensionality of the parameter estimation problem since the measurement equation, Equation (15), is no longer a function of the pulsar distance. We stress that the pulsar terms are dropped only for purposes of Bayesian inference, i.e. from the Kalman filter model that feeds into the nested sampling algorithm. The generated synthetic data that we use to test our method does include the pulsar term in full.

Initially we consider a single noise realisation of the synthetic data for the representative example system described in Table 1. We apply the Kalman filter using Equation (38) in conjunction with nested sampling in order to infer the posterior distributions in each of the parameters. The results are shown in Figure 1 for the seven parameters of  $\theta_{\text{gw}}$ . The histograms shown in the figure are the one-dimensional posteriors for each parameter, marginalized over the set of all other parameters, where the dashed vertical blue lines show the 0.16 and 0.84 quantiles and the solid orange line the true injection value. The two-dimensional contours denote the (0.5, 1, 1.5, 2)-sigma equivalent contours. It is evident that for this representative example with characteristic strain  $h_0 = 5 \times 10^{-15}$  that we are able to estimate all the static parameters of the system

using our method, with the injected value contained within the 90% credible interval for all of the parameters.

For some of the parameters the estimated posterior is approximately symmetric about the true injected value (e.g.  $\Omega$ ). In this case the posterior median and the injected value are approximately coincident. For other parameters (e.g.  $\iota$ ) the distribution is not symmetric about the injected value and the posterior median and the injected value are not coincident. The median value of the posterior for  $\iota$  is shifted by  $\sim 0.35$  rad relative to the injected value, although the injected value does still remain inside the 90% credible interval. Similar effects are seen, albeit with a smaller shift, in other parameters such as  $\delta$  and  $\alpha$ . For a single realisation of the noise it is not clear if this discrepancy is a systematic effect, i.e. a bias in the parameter estimate, or else just a random variation specific to this particular noise realisation. We explore this further in Section ?? and Section 9.1 and show that this is a bias that results from dropping the pulsar terms, similar to that reported in Zhu et al. (2016). Similar results are derived for the the  $3N$  parameters of  $\theta_{\text{psr, reduced}}$  and the parameters are generally recovered unambiguously, but we do not display the resulting posterior distributions here since the inference of  $\theta_{\text{gw}}$  is the main focus of this work. We are able to simultaneously estimate both  $\theta_{\text{gw}}$  and  $\theta_{\text{psr, reduced}}$ .

For each parameter in  $\theta_{\text{gw}}$  we calculate  $W_1$  between every pair of posteriors across our  $10^3$  noise realisations. The results are presented in Figure ?. The figure shows a lower triangular heat-map where each point denotes the  $W_1$  value between a pair of posteriors. Small values of  $W_1$  are magenta moving to red for large values of  $W_1$ . Note that the heat-map colour scaling is not the same for all subplots of the figure. It is evident that  $W_1$  is generally small across all pairs of posteriors. We define  $W_{1, \text{median}}$  as the median value of  $W_1$  across all  $\binom{10^3}{2}$  pairs of noise realisations. The values of  $W_{1, \text{median}}$  for each parameter are quoted in Table ?. As a percentage of the width of the prior space (e.g. for  $\alpha$  the width of the prior is  $2\pi$ ), the median  $W_1$  for each parameter is  $\lesssim 3\%$ . Generally speaking, the method converges to consistent one-dimensional posteriors for each parameter across different noise realisations.

There are a small fraction of the total number of noise realisations which have a large Wasserstein distance from all the other posteriors of all the other noise realisations. These are the strong yellow lines in Figure ?. These correspond exclusively to posteriors which have converged poorly and not found the likelihood global maximum. For instance, the top strong horizontal yellow line in Figure ?, which is identified by an asterisk in the figure, corresponds to a posterior for  $\psi$  which has converged sub-optimally, with an estimated value of  $\psi = 1.01 \pm 0.06$  rad, highly diverged from the true injected value of 2.50 rad. The  $W_1$  value between this posterior and all the posteriors inferred for all other noise realisations is  $\sim 1.7$  radians. We also observe a strong correlation across all noise pairs (Pearson correlation coefficient = 0.96) between the  $W_1$  inferred for  $\alpha$  and  $\psi$ , indicating that noise realisations which corrupt the inference of  $\psi$  similarly affect  $\alpha$ . This is due to multimodal likelihood surface for  $\psi - \alpha$ , shown in Figure ?. We can see that the likelihood surface has a global optima at the approximate location of the injected values ( $\alpha \sim 1.0$ ,  $\psi \sim 2.5$ ) and local optima that corresponds to the location of the example poorly converged posterior discussed previously, with  $\psi = 1.01$  rad. This problem of poor global convergence can be solved in principle by increasing the number of live points  $n_{\text{live}}$  in the nested sampler so as to better cover the likelihood surface and prevent falling into local minima.



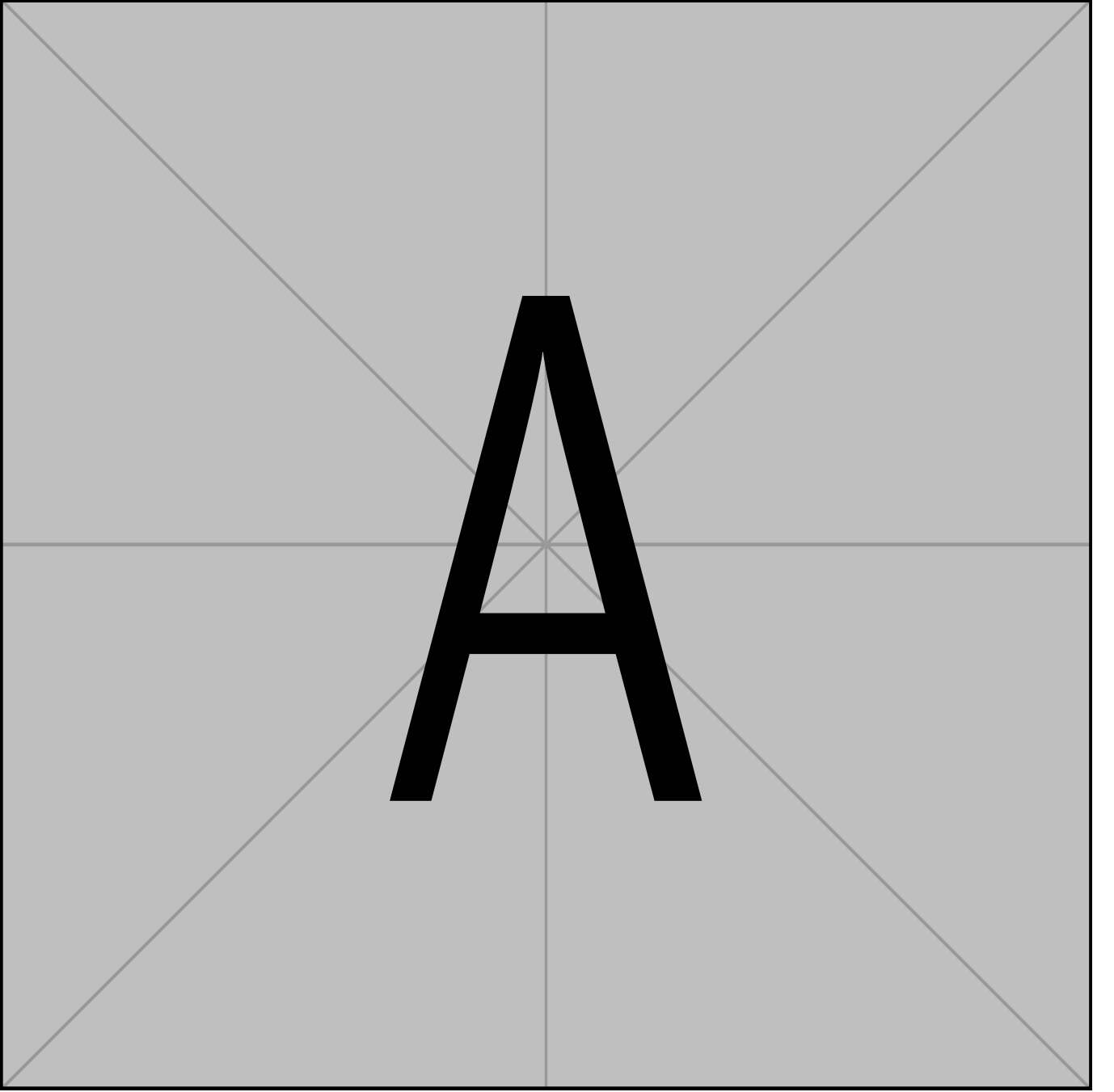
Parameter	Injected Values	Units	Prior	
$\theta_{\text{gw}}$	$\Omega$	$5 \times 10^{-7}$	Hz	$\text{LogUniform}(10^{-9}, 10^{-5})$
	$\alpha$	1.0	radians	$\text{Uniform}(0, 2\pi)$
	$\delta$	1.0	radians	$\text{Uniform}(-\pi/2, \pi/2)$
	$\psi$	2.50	radians	$\text{Uniform}(0, 2\pi)$
	$\Phi_0$	0.20	radians	$\text{Uniform}(0, 2\pi)$
	$h_0$	$5 \times 10^{-15}$	-	$\text{LogUniform}(10^{-15}, 10^{-9})$
	$\iota$	1.0	radians	$\text{Uniform}(0, \pi)$
$\theta_{\text{psr}}$	$f_{\text{em}}^{(n)}(t_1)$	$f_{\text{ATNF}}^{(n)}$	Hz	$\text{Uniform}\left(f_{\text{ATNF}}^{(n)} - 10^3 \epsilon_f^{(n)}, f_{\text{ATNF}}^{(n)} + 10^3 \epsilon_f^{(n)}\right)$
	$\dot{f}_{\text{em}}^{(n)}(t_1)$	$\dot{f}_{\text{ATNF}}^{(n)}$	$\text{s}^{-2}$	$\text{Uniform}\left(\dot{f}_{\text{ATNF}}^{(n)} - 10^3 \epsilon_{\dot{f}}^{(n)}, \dot{f}_{\text{ATNF}}^{(n)} + 10^3 \epsilon_{\dot{f}}^{(n)}\right)$
	$d^{(n)}$	$d_{\text{ATNF}}^{(n)}$	m	-
	$\sigma^{(n)}$	$\sigma_{\text{sc}}^{(n)}$	$\text{s}^{-3/2}$	$\text{LogUniform}(10^{-23}, 10^{-19})$
	$\gamma^{(n)}$	$10^{-13}$	$\text{s}^{-1}$	-

**Table 1.** Summary of injected static parameters used for generating synthetic data in the representative example of Section 7, along with the choice of prior used for Bayesian inference on each parameter. The subscript “ATNF” denotes values which have been obtained from the ATNF pulsar catalogue as described in Section 6.1. The subscript “sc” indicates that the injected value has been calculated using the empirical timing model, Equations (30) and (31) from Shannon & Cordes (2010a). The quantities  $\epsilon_f^{(n)}$  and  $\epsilon_{\dot{f}}^{(n)}$  are the errors in  $f_{\text{em}}^{(n)}(t_1)$  and  $\dot{f}_{\text{em}}^{(n)}(t_1)$  respectively, as quoted in the ATNF catalogue.

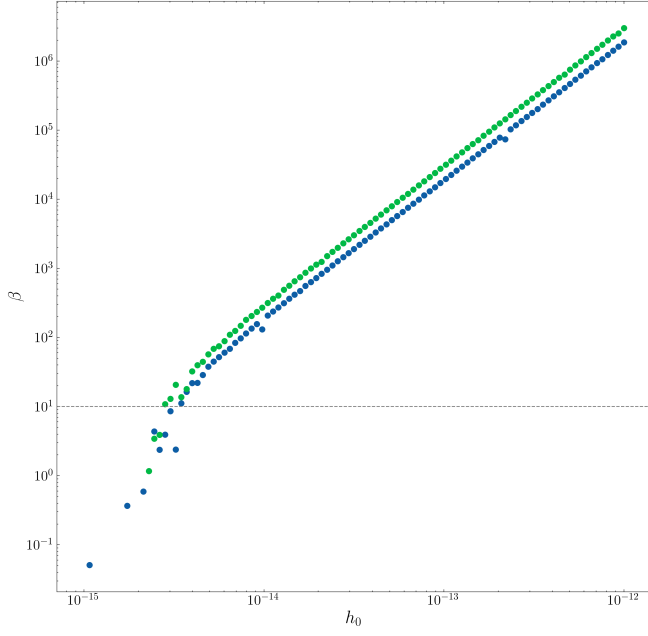
### 7.3 Detection

We frame the problem of claiming a detection of a GW in the noisy data in terms of a Bayesian model selection procedure as described in Section 5.3.  $\mathcal{M}_1$  is the Earth-terms only model, i.e. the state-space model with a Kalman filter using Equation (38). The Bayes factor,  $\beta$ , as defined in Equation (29) is presented in Figure 2 for our representative system, where we now vary the magnitude of the GW strain,  $h_0$ . The noise processes in the synthetic data are identical realisations for each value of  $h_0$ ; the only change in the data is the change in  $h_0$ . We can see that for  $\beta > 10$  there is an approximate log-linear relationship between  $\beta$  and  $h_0$  where the GW source is easily detectable with decisive evidence ( $\beta \sim 100$ ) for  $h_0 \gtrsim 10^{-14}$ . We take  $\beta = 10$  as a tolerance cut-off above which we can accept  $\mathcal{M}_1$  and claim a detection of the GW. For this system, given this realisation of the noise, the minimal detectable strain is  $\sim 3.9 \times 10^{-15}$ . This minimal detectable strain is of course particular to this system and will be influenced in general by e.g.  $T_{\text{obs}}$ ,  $\sigma_{\text{m}}$  as well as the parameters of the GW source such as its location on the sky.

There is an evident sparsity and steep drop-off of  $\beta$  for values of  $h_0 \lesssim 3.9 \times 10^{-15}$ . The steep drop off is due to the two competing models becoming increasingly indistinguishable as the strain signal decreases and the measurement noise starts to dominate. The sparsity is due some of the  $\beta$  value becoming negative at these points. This is a noise artefact of the nested sampling method. For these points the noise in the sampler starts to dominate, the sampler converges sub-optimally, and the hierarchical relationship between  $\mathcal{M}_0$  and  $\mathcal{M}_1$  fails. That is, the requirement that  $\mathcal{Z}(\mathbf{Y}|\mathcal{M}_1) > \mathcal{Z}(\mathbf{Y}|\mathcal{M}_0)$  no longer holds. There is nothing special about these particular points; if one recreates the  $\beta - h_0$  curve by rerunning the nested sampler these sparse locations appear at different values of  $h_0$ , for  $\beta < 10$ . For  $\beta > 10$  the sampling noise is sufficiently small with respect to the signal and the model hierarchy holds. Increasing the number of live points  $n_{\text{live}}$  used by the nested sampler improves the performance at small values of  $h_0$  since as discussed in Section 5.2 the uncertainties in  $\mathcal{Z}$  scale as  $\mathcal{O}(1/\sqrt{n_{\text{live}}})$ . For the results presented in this subsection we set  $n_{\text{live}} = 2000$ .



**Figure 1.** Posterior distribution for the GW source parameters  $\theta_{\text{gw}}$  for the representative system described in Table 1, for a single realisation of the system noise. The vertical orange lines indicate the true injected values. The contours in the 2D histograms denote the (0.5, 1, 1.5, 2)- $\sigma$  levels. The subtitles of the marginalized 1D posteriors denote the posterior median and the 0.16, 0.84 quantiles. We are able to accurately estimate each parameter of interest. We plot the scaled variables  $\Omega_{\text{nHz}} = \Omega \times 10^9$  Hz and  $h_{0, \times 10^{15}} = h_0 \times 10^{15}$ . Note that the  $x$ -axis scaling does not span the total prior space.

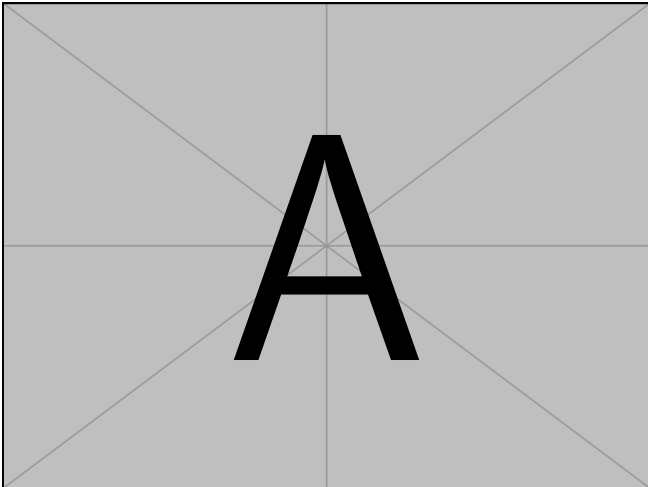


**Figure 2.** Bayes ratio  $\beta$  between the competing models  $\mathcal{M}_1$  (GW present in data) and  $\mathcal{M}_0$  (GW not present in data) at different GW strain magnitudes  $h_0$ , for the representative example summarised in Table 1. The horizontal grey dashed line labels the detection threshold of  $\beta = 10$ . The minimal detectable strain, below which  $\beta < 10$ , is  $\sim 3.9 \times 10^{-15}$ . Gaps in the data below  $\beta = 10$  occur due to noise in the sampler breaking the hierarchical relationship between  $\mathcal{M}_0$  and  $\mathcal{M}_1$ .

## 8 EXPLORING A BROADER PARAMETER SPACE

We have so far focussed on a single representative system, summarised in Table 1. In this section we test the method in different regions of the parameter space, varying the source parameters through astrophysically relevant ranges.

We consider 200 injections for different sets of parameter values where we fix  $h_0 = 5 \times 10^{-15}$ ,  $\iota = 1.0$  and draw the remaining five parameters of  $\theta_{\text{gw}}$  from the prior distributions described in Table 1. We fix  $h_0$  and  $\iota$  in order to maintain an approximately constant SNR across the parameter space. For each simulated injection we can then attempt to recover the posterior distributions for each parameter. To summarise the results across the parameter space we use a parameter-parameter (PP) plot (Cook et al. 2006). A PP plot describes the fraction of the total number of injected parameters which are included within a given credible interval of the estimated posterior with respect to the credible interval itself. In the ideal case the PP plot should be a diagonal line, indicating that  $x$ -percentage of the simulated injections fall within the  $x$ -percentage credible interval. The results are shown in Figure 3. The shaded grey contours enclose the  $1\sigma$ ,  $2\sigma$ , and  $3\sigma$  significance levels, given 200 injections. We can see that only  $\Omega$  falls within the  $3\sigma$  shaded region. The other parameters deviate from the  $y = x$  diagonal. The effect is most pronounced for  $\alpha$  and  $\delta$  with more modest deviations for  $\psi$  and  $\Phi_0$ . The shape of the graph indicates that the posteriors for these parameters are over-constrained; there are fewer injections contained within high value credible intervals than would be expected statistically, and there are more injections contained within the low value credible intervals. This is an result of the bias that we observed in Section ??; the posteriors inferred for these parameters are generally confident, but are systematically biased away from the true injection values. This manifests in the PP plot as posteriors which are overly precise (narrow) to contain the injected value within the appropriate credible interval. An in-depth discussion on the origin of this bias is given in Section 9.1.



**Figure 3.** PP plot for 200 simulated GW sources for a subset of parameters of  $\theta_{\text{gw}}$ , randomly drawn from the prior distributions of Table 1. Each coloured line corresponds to a different parameter of  $\theta_{\text{gw}}$ . The parameters  $h_0$  and  $\iota$  are fixed at  $5 \times 10^{-15}$  and 1.0 rad respectively in order to maintain an approximately constant SNR. The grey shaded contours labels the 1, 2, 3- $\sigma$  confidence intervals. Well estimated posteriors should fall along the diagonal  $y = x$ .  $\Omega$  is generally well-estimated but many other parameters show evidence of being over-constrained due to a modelling bias.

## 9 IDENTIFIABILITY AND BIAS

There are two underlying issues in the preceding tests that we now explore in more detail. The first relates to parameter identifiability, i.e. can we uniquely identify the parameters of the model given the data and the model structure? This question is explored in Section ???. The second relates to a bias in the parameter estimates that results from dropping the pulsar terms from the Kalman filter measurement equation as described in Section 7.2. This is the focus of Section 9.1. In order to elucidate these questions without confusion from the measurement noise, we now work in the high SNR regime and set  $h_0 = 10^{-12}$  (c.f. Figure 2) for our representative example system (Table 1).<sup>4</sup>

### 9.1 Bias by dropping pulsar terms

Evident in the preceding discussion is the fact that for some of the  $\theta_{\text{gw}}$  parameters the inferred value (e.g. the median of the one-dimensional marginalised posterior) is biased away from the true injection value. In Figure ?? we show the resulting posteriors for  $\theta_{\text{gw}}$ , analogous to Figure 1, but now for the high-SNR system with  $h_0 = 10^{-12}$ . We can see that for all parameters, with the exception of  $\Omega$  and  $h_0$ , the injected value lies outside the 90% credible interval, to the extent that it is not visible on the scales over which we have plotted the posteriors. This effect is most severe for  $\iota$  with a bias of  $\sim 0.3$  rad but is also present to a lesser extent in other static parameters such as  $\delta$ ,  $\alpha$  and  $\psi$ , and additionally to an even smaller extent in  $\Phi_0$ . Similar results are obtained for other noise realisations. Whilst  $h_0$  does not exhibit any noticeable bias for this particular noise realisation, in general across multiple noise realisations it does exhibit a correlated bias with  $\iota$ . In this Section we demonstrate how this bias results from the dropping of the pulsar terms described in Section 7.2.

In Figure ?? we plot the variation in the log-likelihood returned by the Kalman filter, Equation (26), as we vary  $\iota$  across the prior space, holding all other parameters constant at the true injected value of the representative example system, Table 1. We do this for 10 separate realisations of the system noise. We consider three separate situations: (i) The solid blue line is the likelihood curve with just the Earth-terms included (i.e. using Equation 38 in the Kalman filter), (ii) the solid orange line is the likelihood curve with the additional pulsar terms included (i.e. using Equation 15), (iii) the solid green line is the solution where we drop the pulsar terms from both the synthetic data and the measurement equation. We have not used nested sampling for situations (ii) or (iii) in this work. However, we are able to run the Kalman filter either using the pulsar terms in the model or else dropping them from the synthetic data consider how the resulting likelihood varies for different parameter values. The dashed lines show the location of the likelihood maxima for each of the likelihood solutions; finding the location of the maxima across all static parameters is effectively the goal of any likelihood based inference method, such as nested sampling.

The key observation from Figure ?? is that for the Earth-terms only solution the location of the likelihood optima and the location of the injected value are not the same. Consequently Bayesian likelihood inference techniques settle in optima that do not necessarily align

<sup>4</sup> TK: This discussion on bias and identifiability is perhaps less relevant than it was when our representative example had a 'high' strain ( $10^{-12}$ ). However it IS still a feature/bug of our approach and I think it is worth discussing in the main body rather than relegating to an appendix. Opinions welcome...



with the true injection value. This is the origin of the bias in the one-dimensional marginalised posteriors that we observe. For the Earth-terms solution the optima of the likelihood curve lies around  $\iota \sim 0.7$  radians, whereas the injected value was  $\iota = 1.0$  radians. This optima is comparable to the median of the posterior inferred for  $\iota$  in Figure ??.

In contrast to the Earth-terms solution, for the pulsar-terms solution the optima of the likelihood curve now aligns with the true injected value. The corollary is also true; if we drop the pulsar terms from the synthetic data and use the Earth terms equation for inference then the optima is also correctly located. This evidences that it is the effect of dropping the pulsar terms from the measurement equation which causes this bias in the inferred posteriors. Similar results are observed for the other parameters which exhibit a bias in the nested sampling inference.

Why does dropping the pulsar terms bias these parameters in particular, whilst leaving others unaffected? For instance, why is the bias induced in  $\iota$  so strong, whereas  $\psi$  shows only a weak bias and  $\Omega$  shows no bias at all? When we generate synthetic data including the pulsar term, the state frequency of an individual pulsar  $f_p$  is effectively modulated by the interference of two equal-amplitude, phase shifted waves (c.f. Equation 15). Explicitly, we have the “Earth-terms wave” with a general form  $y_{\text{Earth}} = A \cos \xi$  for amplitude  $A$  and phase  $\xi$  and the “pulsar-terms wave” with a general form  $y_{\text{pulsar}} = A \cos(\xi + \kappa)$  for phase offset  $\kappa$ . The net interference wave that modulates  $f_p$  is

$$y_{\text{Earth}} - y_{\text{pulsar}} = 2A \sin\left(\frac{\kappa}{2}\right) \sin\left(\xi + \frac{\kappa}{2}\right). \quad (40)$$

The additional modulations from  $y_{\text{pulsar}}$  influence the amplitude of the resulting interference wave, but do not affect its frequency. When we model this data using the Kalman filter without the pulsar terms, the additional amplitude modulations from  $y_{\text{pulsar}}$  must be described by the single wave model,  $y_{\text{Earth}}$ . As a result the  $y_{\text{Earth}}$  model must try to compensate for these amplitude modulations by adjusting the parameters which govern the amplitude of the single wave. Consequently the parameters which are most affected are those which determine the amplitude of the GW. Terms which influence the amplitude strongly, such as  $h_0$  and  $\iota$ , are strongly affected and exhibit a strong bias. Terms which influence the amplitude weakly such as  $\psi$  exhibit a weaker bias. The wave amplitude is independent of the frequency and so  $\Omega$  exhibits no bias. Since a shift in phase affects the instantaneous amplitude of the wave,  $\Phi_0$  also exhibits a small bias, despite  $A$  not being a function of  $\Phi_0$ . This hierarchy of biases for each parameter is also what we observe in Figure 3, where  $\alpha$  and  $\delta$  are strongly over-constrained whilst  $\Omega$  is well-estimated.

In order to ascertain how important the bias is for PTA continuous wave searches requires a thorough exploration of the astrophysical parameter space which we do not carry out in this work. The magnitude of the bias will vary for different points in the parameter space, as well as for different PTA configurations. If the bias is sufficiently modest then for quiet continuous wave sources with low SNR the uncertainty in the one-dimensional marginalised posterior may dominate and the bias may not be important for searches. This is what we observe for our synthetic data (e.g. Figure ??); the decreased SNR manifests as a broader marginalised posterior which generally overwhelms the shift due to the bias. Conversely if the bias is sufficiently strong, or the source sufficiently loud with low uncertainties, then the bias may come to systematically affect the

inferred parameters.

We have shown in this section how the dropping of the pulsar terms results in a bias in the inferred parameters, not just in the sky localisation parameters (e.g. Zhu et al. 2016; Chen & Wang 2022) but additionally all parameters which influence the wave amplitude. Whilst the bias is generally small and does not prevent the recovery of accurate posteriors for our example system, it is present and may need to be accounted for for analysis in other parts of the parameter space. This bias is not particular to our method but a shared feature across likelihood based methods that do not include the pulsar terms in the inference model. The inclusion of the pulsar terms in the methods presented in this work in order to alleviate the bias will be a key goal of a future work.

## 10 DISCUSSION

We have shown how our complementary approach to PTA data analysis using a Kalman filter can successfully recover the underlying parameters of the model and calculate Bayes factors in order to perform Bayesian model selection. There are a number of potential extensions of this work which we now discuss.

The natural first extension is to implement the method whilst retaining the pulsar terms of Equation 15. Whilst PTA searches commonly drop the pulsar terms as standard practice, we have shown that this leads to biases in the inferred one-dimensional marginalized posteriors for the model parameters. Biases in sky localisation for synthetic datasets as a result of dropping the pulsar terms are also observed by Zhu et al. (2016) and Chen & Wang (2022). Alternatively, if the method continues to be used with solely the Earth terms then it would be highly desirable to make systematic, quantitative estimates of the incurred biases in the model parameters across an astrophysically representative parameter space.

We have considered only a specific configuration of pulsars composing our synthetic PTA - the same pulsars that make up the 12.5-year NANOGrav (Section 6.1). Whilst our method is generally independent of the specific choice of pulsars, it would be of interest to compare the performance of the method using different pulsars configurations and different PTAs such as PPTA and the EPTA. Indeed, since adding pulsars to PTAs increases the computational demands of the data analysis it may be advantageous to optimally select a subset of the total available pulsars which may be able to provide comparable performance to a larger PTA (Speri et al. 2023).

Throughout this work we have assumed that all pulsars are observed at the same time, with a constant time sampling. In practice this assumption does not hold and different pulsars will be observed with different cadences at different times. Extending this method to non-constant time sampling is straightforward and the performance of the method should be evaluated in this regime.

We have taken our GW source to be monochromatic, i.e.  $\Omega$  is a fixed parameter with no time dependence. Whilst this is an astronomically well-justified assumption (see discussion in Section 2.2.1), and appropriate for this initial methods work, it would also be of interest to extend the state-space construction to enable the GW frequency to evolve in time. The GW source is monochromatic over the decadal PTA observation period, but the timescale set by the Earth-pulsar distance is comparable to the SMBHB evolution

timescale (Sesana & Vecchio 2010). For  $\Delta f_{\text{gw}} < 1/T_{\text{obs}}$  (see Equation (6)) the GW source frequency evolution manifests as an incoherent source of noise in the pulsar terms, whilst for  $\Delta f_{\text{gw}} > 1/T_{\text{obs}}$  the pulsar terms can affect the phase coherency of the Earth terms (Perrodin & Sesana 2018). Careful consideration of the evolution of the GW frequency will be important as we look to extend our method to include the pulsar terms in the inference model.

We have also assumed that there is only a single GW source that influences the received pulsar frequencies. However, it may be possible to resolve multiple continuous GW sources concurrently (Babak & Sesana 2012b). Our method naturally extends to detecting multiple GW sources simultaneously; Equation (15) is straightforwardly modified to become a linear superposition of multiple GWs. The stochastic background itself is the incoherent summation of many individual GW sources. It therefore stands to reason that it may be possible to extend this method to also search for a stochastic GW background, providing a complementary approach to detect the stochastic background without requiring a cross-correlation of the pulsar residuals to uncover the Hellings-Downs curve.

**TK:** Lets also pre-empt reviewer and make it explicit why we are using frequencies rather than TOAs. Need to chat with group to get this clear in my head.

## 11 CONCLUSION

In this paper we demonstrate a new method for the detection and parameter estimation of GWs from individual, monochromatic SMBHBs. We track the evolution of the intrinsic pulsar timing noise explicitly via a state-space method, rather than subtracting the ensemble-averaged statistics. That is, we disentangle statistically the specific time-ordered realisation of the timing noise from the GW induced modulations. This enables the inference of GW parameters conditioned on the specific observed realisation of the noisy data. We introduce the Kalman filter in order to track the intrinsic state evolution of the pulsar and combine this with a Bayesian sampling framework to estimate the posterior distributions of each parameter of the system, as well as the associated evidence (marginalised likelihood) of the model. We test our method on synthetic data. We initially focus on a single, astrophysically representative, SMBHB GW source detected with the 12.5-year NANOGrav pulsars observed over a 10 year timespan. The parameters of the model are accurately recovered and the minimal detectable strain is estimated to be  $h_0 \sim 4 \times 10^{-15}$  for  $\iota = 1.0$  rad. The method is then further tested over 1000 noise realisations. Consistent posterior estimates of the parameters are obtained for the majority of noise realisations, with the median Wasserstein distance across all noise realisations  $W_{1,\text{median}}$  generally small compared to the prior space ( $\lesssim 3\%$ ). For a minority of noise realisations the nested sampler converges poorly and settles into a local likelihood optima for the posteriors estimates of  $\psi$  and  $\alpha$ . This problem can be solved in general by increasing the number of live points  $n_{\text{live}}$  used by the sampler. Exploration of a broader parameter space via 200 randomly sampled parameter vectors highlights biases in some parameters as a result of dropping the pulsar terms from the inference model. The cause of these biases is examined for the case of high SNR, along with the structural weak-identifiability between  $\iota$  and  $h_0$ . The extension of this method to include the pulsar term is a key challenge for a future work.

## APPENDIX A: KALMAN RECURSION EQUATIONS

The linear Kalman filter operates on temporally discrete, noisy measurements  $\mathbf{Y}_k$ , which are related to a set of unobservable discrete system states  $\mathbf{X}_k$ , via a linear transformation:

$$\mathbf{Y}_k = \mathbf{H}_k \mathbf{X}_k + \mathbf{v}_k, \quad (\text{A1})$$

where  $\mathbf{H}_k$  is the measurement matrix or observation model,  $\mathbf{v}_k$  is a zero-mean Gaussian measurement noise,  $\mathcal{N} \sim (0, \mathbf{R}_k)$  with covariance, and the subscript  $k$  labels the time-step. The Kalman filter takes the underlying states to evolve according to

$$\mathbf{X}_k = \mathbf{F}_k \mathbf{X}_{k-1} + \mathbf{G}_k \mathbf{u}_k + \mathbf{w}_k, \quad (\text{A2})$$

where  $\mathbf{F}$  is the system dynamics matrix,  $\mathbf{G}$  the control matrix.  $\mathbf{u}$  the control vector, and  $\mathbf{w}_k$  is a zero-mean Gaussian process noise,  $\mathcal{N} \sim (0, \mathbf{Q}_k)$ .

The Kalman filter is a recursive estimator with two distinct stages; a “predict” stage and an “update” stage. The predict stage predicts the estimate of the state at discrete step  $k$ , given the state estimates from step  $k-1$ . Specifically, the predict step proceeds as,

$$\hat{\mathbf{X}}_{k|k-1} = \mathbf{F}_k \hat{\mathbf{X}}_{k-1|k-1} + \mathbf{G}_k \mathbf{u}_k \quad (\text{A3})$$

$$\hat{\mathbf{P}}_{k|k-1} = \mathbf{F}_k \hat{\mathbf{P}}_{k-1|k-1} \mathbf{F}_k^\top + \mathbf{Q}_k \quad (\text{A4})$$

Note that the predict stage is independent of the measurements. The measurement  $\mathbf{Y}_k$  is included to update the prediction during the update stage as follows:

$$\boldsymbol{\epsilon}_k = \mathbf{Y}_k - \mathbf{H}_k \hat{\mathbf{X}}_{k|k-1} \quad (\text{A5})$$

$$\mathbf{S}_k = \mathbf{H}_k \hat{\mathbf{P}}_{k|k-1} \mathbf{H}_k^\top + \mathbf{R}_k \quad (\text{A6})$$

$$\mathbf{K}_k = \hat{\mathbf{P}}_{k|k-1} \mathbf{H}_k^\top \mathbf{S}_k^{-1} \quad (\text{A7})$$

$$\hat{\mathbf{X}}_{k|k} = \hat{\mathbf{X}}_{k|k-1} + \mathbf{K}_k \boldsymbol{\epsilon}_k \quad (\text{A8})$$

$$\hat{\mathbf{P}}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \hat{\mathbf{P}}_{k|k-1} \quad (\text{A9})$$

Equation (A7) defines the “Kalman gain”  $\mathbf{K}_k$  which is defined so as to minimise the mean squared error in the state estimate, i.e.  $\mathbf{K}_k = \text{argmin} [\mathbf{E}(\mathbf{X}_k - \hat{\mathbf{X}}_k)]$ . For a full review of the Kalman filter, including its derivation, we refer the reader to Gelb et al. (1974) and Zarchan & Musoff (2000).

We map our continuous-time state-space model described in Section 2 onto the discrete-time Kalman filter structure of Equations (A1), (A2) as follows. We identify  $\mathbf{X}$  with a vector of length  $N$  composed of the intrinsic pulsar frequency states, i.e

$$\mathbf{X} = (f_p^{(1)}, f_p^{(2)}, \dots, f_p^{(N)}) \quad (\text{A10})$$

Similarly,

$$\mathbf{Y} = (f_m^{(1)}, f_m^{(2)}, \dots, f_m^{(N)}) \quad (\text{A11})$$

The states evolve according the continuous dynamical equation (c.f. Equation (2))

$$d\mathbf{X} = \mathbf{A}\mathbf{X}dt + \mathbf{C}(t)dt + \boldsymbol{\Sigma}d\mathbf{B}(t) \quad (\text{A12})$$

where  $\mathbf{A}$  is a diagonal  $N \times N$  square matrix

$$\mathbf{A} = \text{diag} \left( \gamma^{(1)}, \gamma^{(2)}, \dots, \gamma^{(N)} \right), \quad (\text{A13})$$

and  $\mathbf{C}(t)$  a time dependent vector with  $n$ -th component

$$C^{(n)} = \gamma^{(n)} \left( f_{\text{em}}^{(n)}(t_1) + \dot{f}_{\text{em}}(t_1)^{(n)} t \right) + \dot{f}_{\text{em}}(t_1)^{(n)}. \quad (\text{A14})$$

The  $N \times N$  square matrix  $\Sigma$  governs the magnitude of the increments of Brownian motion (Wiener process)  $d\mathbf{B}(t)$ , where

$$\Sigma = \text{diag} \left( \sigma^{(1)}, \sigma^{(2)}, \dots, \sigma^{(N)} \right) \quad (\text{A15})$$

The evolution of the state is an Ornstein-Uhlenbeck process, described by a Langevin equation, Equation (A12), which has a general solution given by (Gardiner 2009),

$$\mathbf{X}(t) = e^{\mathbf{A}t} \mathbf{X}(0) + \int_0^t e^{\mathbf{A}(t-t')} \mathbf{C}(t') dt' + \int_0^t e^{\mathbf{A}(t-t')} \Sigma d\mathbf{B}(t') \quad (\text{A16})$$

From Equation (A16) we can construct the discrete, recursive solution for  $\mathbf{X}(t_k) = \mathbf{X}_k$  in the form of Equation (A2) where

$$\mathbf{F}_k = e^{\mathbf{A}(t_{k+1}-t_k)} \quad (\text{A17})$$

$$\mathbf{G}_k = \int_{t_i}^{t_{i+1}} e^{\mathbf{A}(t_{k+1}-t')} \mathbf{C}(t') dt' \quad (\text{A18})$$

$$\mathbf{w}_k = \int_{t_k}^{t_{k+1}} e^{\mathbf{A}(t_{k+1}-t')} \Sigma d\mathbf{B}(t') \quad (\text{A19})$$

For our system, the state, control and process noise covariance matrices then take the form:

$$\mathbf{F}_k = \text{diag} \left( e^{-\gamma^{(1)} \Delta t}, e^{-\gamma^{(2)} \Delta t}, \dots, e^{-\gamma^{(N)} \Delta t} \right) \quad (\text{A20})$$

$$\mathbf{G}_k = \left( G_k^{(1)}, G_k^{(2)}, \dots, G_k^{(N)} \right) \quad (\text{A21})$$

with

$$G_k^{(n)} = f_{\text{em}}^{(n)}(t_1) + \dot{f}_{\text{em}}^{(n)}(t_1) [\Delta t + t_k] - e^{-\gamma \Delta t} \left[ f_{\text{em}}^{(n)}(t_1) + \dot{f}_{\text{em}}^{(n)}(t_1) t_k \right] \quad (\text{A22})$$

and

$$\mathbf{Q}_k \delta_{kj} = \langle \boldsymbol{\eta}_k \boldsymbol{\eta}_j^\top \rangle = \text{diag} \left( Q^{(1)}, Q^{(2)}, \dots, Q^{(N)} \right) \quad (\text{A23})$$

for

$$Q^{(n)} = \frac{[\sigma^{(n)}]^2}{2\gamma^{(n)}} \left( e^{-2\gamma^{(n)} \Delta t} - 1 \right) \quad (\text{A24})$$

where  $\Delta t$  is the constant time sampling interval  $= t_{k+1} - t_k$ .

The remaining component matrices of the Kalman filter that we have not yet specified are the measurement matrix  $\mathbf{H}_k$  and the measurement covariance matrix  $\mathbf{R}_k$ . These are defined straightforwardly from Equation (15). Specifically,  $\mathbf{H}_k$  is a diagonal matrix where the  $n$ -th component of the diagonal is given by  $g^{(n)}(t_k)$  from Equation (18). The measurement covariance  $\mathbf{R}_k = E[\mathbf{v}\mathbf{v}^\top] = \sigma_{\text{m}}^2$ .

## APPENDIX B: FIDUCIAL PULSAR VALUES

**TK:** Do we really need this? The argument is for reproducibility, but we have a nice github repo that holds all the data we actually used, jupyter notebooks that show how the calculations are done. Perhaps it is sufficient to point the reader to the repo...

## APPENDIX C: IDENTIFIABILITY

**TK TBD.** An example calculation for a simple system using the methods of e.g. (Karlsson et al. 2012) (Sedoglavic 2002)

## APPENDIX D: WASSERSTEIN DISTANCE

The Wasserstein distance is a metric that defines a distance between two probability distributions  $\mu(x)$  and  $\nu(x)$ . It has an intuitive interpretation as the lowest total cost with which one can move probability mass from  $\mu$  to  $\nu$ , with respect to a cost function  $c(x, y)$ . For this reason it is sometimes known as the "Earth mover's distance". The  $p$ -th order Wasserstein distance between two distributions is

$$W_p(\mu, \nu) = \left( \inf_{\gamma \in \Gamma(\mu, \nu)} \int c(x, y)^p d\gamma(x, y) \right)^{1/p}, \quad (\text{D1})$$

where  $p \in [1, \infty)$  and  $\Gamma(\mu, \nu)$  is the set of all joint probability distributions for  $(x, y)$  that have marginals  $\mu$  and  $\nu$ , i.e.  $\Gamma(\mu, \nu)$  is the set of couplings of  $\mu$  and  $\nu$ .

The cost function can be freely chosen for the nature of the problem, but is commonly taken to be the absolute value function,  $c(x, y) = |x - y|$ , and we adopt the same cost function in this work. Given this choice of  $c(x, y)$ , for an empirical distribution  $\hat{\mu}_n$  composed of  $n$  i.i.d samples drawn from the underlying probability distribution  $\mu$ , the solution to the discrete form of Equation (D3) can be computed generally by the Hungarian algorithm (Kuhn 1955) in polynomial time  $O(n^3)$  (Villani 2009). However, for  $\mu$  defined on  $\mathbb{R}^d$  it is well-known (Dudley 1969) that the Wasserstein distance suffers from the curse of dimensionality for dimensions  $d \geq 3$ :

$$E[W_1(\hat{\mu}_n, \mu)] \asymp n^{-1/d}. \quad (\text{D2})$$

In order to circumvent this issue, in this work we calculate the Wasserstein distance exclusively between the one-dimensional marginalized posteriors, i.e. setting  $d = 1$ . In this case Equation (D3) has a convenient, easily computed, closed form:

$$W_p(\mu, \nu) = \left( \int_0^1 F_\mu^{-1}(z) - F_\nu^{-1}(z) dz \right)^{1/p} \quad (\text{D3})$$

where  $F_\mu(z)$  is the cumulative distribution function of  $\mu$ .

The use of the Wasserstein distance over more common probability measures such as the Kullback-Leibler divergence or the Kolmogorov-Smirnoff test has a number of advantages. Firstly, it is intuitive, being the minimal amount of work required to transform one distribution into another. It also inherits the units of the underlying distributions, providing a convenient heuristic; if  $W_1 = 1$  between a pair of distributions with units e.g. radians then we can interpret this as the cost to transform between the two distributions is 1 radian. Further interpretability is provided by the Monge-Kantorovich duality, Equation (??); the difference in the expected value of two random variables drawn from the two distributions is less than 1

radian. Secondly it is a versatile, non-parametric measure that requires no specific properties of the probability distributions that are being compared such as e.g. Gaussianity. Any two distributions can be compared, regardless of whether they are continuous, discrete, singular or defined on different metric spaces. Thirdly, it is a true distance metric and as such satisfies desirable properties of a measure of distance such as symmetry and the triangle inequality. It also respects the geometry of the underlying metric space over which the distributions are defined.

## REFERENCES

- Agazie G., et al., 2023a, *ApJ*, **951**, L8  
 Agazie G., et al., 2023b, *ApJ*, **951**, L9  
 Alam M. F., et al., 2021, *ApJS*, **252**, 5  
 Allen B., 1997, in Marck J.-A., Lasota J.-P., eds, *Relativistic Gravitation and Gravitational Radiation*. pp 373–417 ([arXiv:gr-qc/9604033](https://arxiv.org/abs/gr-qc/9604033)), [doi:10.48550/arXiv.gr-qc/9604033](https://doi.org/10.48550/arXiv.gr-qc/9604033)  
 Antonelli M., Basu A., Haskell B., 2023, *MNRAS*, **520**, 2813  
 Antoniadis J., et al., 2022, *MNRAS*, **510**, 4873  
 Antoniadis J., et al., 2023a, *arXiv e-prints*, p. [arXiv:2306.16214](https://arxiv.org/abs/2306.16214)  
 Antoniadis J., et al., 2023b, *arXiv e-prints*, p. [arXiv:2306.16224](https://arxiv.org/abs/2306.16224)  
 Antoniadis J., et al., 2023c, *arXiv e-prints*, p. [arXiv:2306.16226](https://arxiv.org/abs/2306.16226)  
 Arzoumanian Z., et al., 2018a, *ApJS*, **235**, 37  
 Arzoumanian Z., et al., 2018b, *ApJ*, **859**, 47  
 Arzoumanian Z., et al., 2020a, *ApJ*, **900**, 102  
 Arzoumanian Z., et al., 2020b, *ApJ*, **905**, L34  
 Arzoumanian Z., et al., 2023, *arXiv e-prints*, p. [arXiv:2301.03608](https://arxiv.org/abs/2301.03608)  
 Ashton G., Talbot C., 2021, *MNRAS*, **507**, 2037  
 Ashton G., et al., 2022, *Nature Reviews Methods Primers*, **2**, 39  
 Babak S., Sesana A., 2012a, *Phys. Rev. D*, **85**, 044034  
 Babak S., Sesana A., 2012b, *Phys. Rev. D*, **85**, 044034  
 Babak S., et al., 2016, *MNRAS*, **455**, 1665  
 Bailes M., et al., 2020, *Publications of the Astronomical Society of Australia*, **37**, e028  
 Baym G., Pethick C., Pines D., Ruderman M., 1969, *Nature*, **224**, 872  
 Bhagwat S., De Luca V., Franciolini G., Pani P., Riotto A., 2021, *J. Cosmology Astropart. Phys.*, **2021**, 037  
 Brook P. R., Karastergiou A., Buchner S., Roberts S. J., Keith M. J., Johnston S., Shannon R. M., 2013, *The Astrophysical Journal Letters*, **780**, L31  
 Buchner J., 2021a, *arXiv e-prints*, p. [arXiv:2101.09675](https://arxiv.org/abs/2101.09675)  
 Buchner J., 2021b, *The Journal of Open Source Software*, **6**, 3001  
 Burke-Spolaor S., et al., 2019, *A&ARv*, **27**, 5  
 Caballero R. N., et al., 2016, *MNRAS*, **457**, 4421  
 Charisi M., Taylor S. R., Witt C. A., Runnoe J., 2023, *arXiv e-prints*, p. [arXiv:2304.03786](https://arxiv.org/abs/2304.03786)  
 Chen J.-W., Wang Y., 2022, *ApJ*, **929**, 168  
 Christensen N., 2019, *Reports on Progress in Physics*, **82**, 016903  
 Cook S. R., Gelman A., Rubin D. B., 2006, *Journal of Computational and Graphical Statistics*, **15**, 675  
 Cordes J. M., Greenstein G., 1981, *ApJ*, **245**, 1060  
 Cordes J. M., Lazio T. J. W., 2002, *arXiv e-prints*, pp astro-ph/0207156  
 Cordes J. M., Shannon R. M., 2010, *arXiv e-prints*, p. [arXiv:1010.3785](https://arxiv.org/abs/1010.3785)  
 D'Alessandro F., McCulloch P. M., Hamilton P. A., Deshpande A. A., 1995, *MNRAS*, **277**, 1033  
 Desvignes G., et al., 2016, *Monthly Notices of the Royal Astronomical Society*, **458**, 3341  
 Dudley R. M., 1969, *The Annals of Mathematical Statistics*, **40**, 40  
 Dunn L., et al., 2022, *MNRAS*, **512**, 1469  
 Dunn L., Melatos A., Espinoza C. M., Antonopoulou D., Dodson R., 2023, *MNRAS*, **522**, 5469  
 Ellis J. A., 2013, *Classical and Quantum Gravity*, **30**, 224004  
 Ellis J. A., Cornish N. J., 2016, *Phys. Rev. D*, **93**, 084048  
 Ellis J. A., Siemens X., Creighton J. D. E., 2012, *ApJ*, **756**, 175  
 Espinoza C. M., Antonopoulou D., Dodson R., Stepanova M., Scherer A., 2021, *A&A*, **647**, A25  
 Feroz F., Hobson M. P., 2008, *MNRAS*, **384**, 449  
 Feroz F., Hobson M. P., Bridges M., 2009, *Monthly Notices of the Royal Astronomical Society*, **398**, 1601  
 Gardiner C., 2009, *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer Series in Synergetics, Springer Berlin Heidelberg, <https://books.google.com.au/books?id=otg3PQAACAAJ>  
 Gelb A., Kasper J. F., Nash R. A., Price C. F., Sutherland A. A., eds, 1974, *Applied Optimal Estimation*. MIT Press, Cambridge, MA  
 Goldreich P., Julian W. H., 1969, *ApJ*, **157**, 869  
 Goldstein J. M., Veitch J., Sesana A., Vecchio A., 2018, *MNRAS*, **477**, 5447  
 Goncharov B., et al., 2021, *MNRAS*, **502**, 478  
 Greenstein G., 1970, *Nature*, **227**, 791  
 Gügercinoğlu E., Alpar M. A., 2017, *MNRAS*, **471**, 4827  
 Handley W. J., Hobson M. P., Lasenby A. N., 2015, *MNRAS*, **450**, L61  
 Hellings R. W., Downs G. S., 1983, *ApJ*, **265**, L39  
 Hobbs G. B., Edwards R. T., Manchester R. N., 2006, *Monthly Notices of the Royal Astronomical Society*, **369**, 655  
 Hobbs G., Lyne A. G., Kramer M., 2010, *MNRAS*, **402**, 1027  
 Hobbs G., Dai S., Manchester R. N., Shannon R. M., Kerr M., Lee K.-J., Xu R.-X., 2019, *Research in Astronomy and Astrophysics*, **19**, 020  
 Jaffe A. H., Backer D. C., 2003, *The Astrophysical Journal*, **583**, 616  
 Jankowski F., et al., 2018, *Monthly Notices of the Royal Astronomical Society*, **484**, 3691  
 Jenet F. A., Lommen A., Larson S. L., Wen L., 2004, *ApJ*, **606**, 799  
 Johnston S., Galloway D., 1999, *Monthly Notices of the Royal Astronomical Society*, **306**, L50  
 Jones P. B., 1990, *MNRAS*, **246**, 364  
 Kalman R. E., 1960, *Journal of Basic Engineering*, **82**, 35  
 Karlsson J., Anguelova M., Jirstrand M., 2012, *IFAC Proceedings Volumes*, **45**, 941  
 Kerr M., Hobbs G., Johnston S., Shannon R. M., 2015, *Monthly Notices of the Royal Astronomical Society*, **455**, 1845  
 Kimpson 2023, *Geophys. Res. Lett.*, **50**, e2023GL103880  
 Kramer M., Lyne A. G., O'Brien J. T., Jordan C. A., Lorimer D. R., 2006, *Science*, **312**, 549  
 Kuhn H. W., 1955, *Naval Research Logistics Quarterly*, **2**, 83  
 Lam M. T., et al., 2019, *ApJ*, **872**, 193  
 Lasky P. D., Melatos A., Ravi V., Hobbs G., 2015, *MNRAS*, **449**, 3293  
 Lee K. J., Wex N., Kramer M., Stappers B. W., Bassa C. G., Janssen G. H., Karuppusamy R., Smits R., 2011, *MNRAS*, **414**, 3251  
 Lentati L., et al., 2015, *MNRAS*, **453**, 2576  
 Lower M. E., et al., 2021, *MNRAS*, **508**, 3251  
 Lyne A., Hobbs G., Kramer M., Stairs I., Stappers B., 2010, *Science*, **329**, 408  
 Maggiore M., 2018, *Gravitational Waves: Volume 2: Astrophysics and Cosmology*. Oxford University Press, [doi:10.1093/oso/9780198570899.001.0001](https://doi.org/10.1093/oso/9780198570899.001.0001), <https://doi.org/10.1093/oso/9780198570899.001.0001>  
 Manchester R. N., Hobbs G. B., Teoh A., Hobbs M., 2005, *AJ*, **129**, 1993  
 McWilliams S. T., Ostriker J. P., Pretorius F., 2014, *The Astrophysical Journal*, **789**, 156  
 Melatos A., Link B., 2014, *MNRAS*, **437**, 21  
 Melatos A., Peralta C., Wyithe J. S. B., 2008, *ApJ*, **672**, 1103  
 Melatos A., Dunn L. M., Suvorova S., Moran W., Evans R. J., 2020, *ApJ*, **896**, 78  
 Melatos A., O'Neill N. J., Meyers P. M., O'Leary J., 2023, *ApJ*, **944**, 64  
 Meyers P. M., Melatos A., O'Neill N. J., 2021a, *MNRAS*, **502**, 3113  
 Meyers P. M., O'Neill N. J., Melatos A., Evans R. J., 2021b, *MNRAS*, **506**, 3349  
 Mukherjee P., Parkinson D., Liddle A. R., 2006, *ApJ*, **638**, L51  
 Pártay L. B., Bartók A. P., Csányi G., 2009, *arXiv e-prints*, p. [arXiv:0906.3544](https://arxiv.org/abs/0906.3544)  
 Parthasarathy A., et al., 2021, *MNRAS*, **502**, 407  
 Pennucci T. T., Collaboration T. N., 2020, *The NANOGrav 12.5-year Wide-band Data Set (version 12yv4)*, [doi:10.5281/zenodo.4312887](https://doi.org/10.5281/zenodo.4312887), <https://doi.org/10.5281/zenodo.4312887>  
 Peralta C., Melatos A., Giacobello M., Ooi A., 2006, *ApJ*, **651**, 1079  
 Perera B. B. P., et al., 2019, *MNRAS*, **490**, 4666  
 Perrodin D., Sesana A., 2018, in Rezzolla L., Pizzochero P., Jones D. I., Rea



- N., Vidiña I., eds, *Astrophysics and Space Science Library* Vol. 457, *Astrophysics and Space Science Library*. p. 95 ([arXiv:1709.02816](#)), doi:10.1007/978-3-319-97616-7\_3
- Petiteau A., Babak S., Sesana A., de Araujo M., 2013, *Phys. Rev. D*, 87, 064036
- Pitkin M., 2018, *Journal of Open Source Software*, 3, 538
- Price S., Link B., Shore S. N., Nice D. J., 2012, *MNRAS*, 426, 2507
- Rajagopal M., Romani R. W., 1995, *ApJ*, 446, 543
- Ravi V., Wyithe J. S. B., Hobbs G., Shannon R. M., Manchester R. N., Yardley D. R. B., Keith M. J., 2012, *ApJ*, 761, 84
- Ravi V., Wyithe J. S. B., Shannon R. M., Hobbs G., 2015, *MNRAS*, 447, 2772
- Reardon D. J., et al., 2023, *ApJ*, 951, L6
- Renzini A. I., Goncharov B., Jenkins A. C., Meyers P. M., 2022, *Galaxies*, 10, 34
- Rosado P. A., Sesana A., Gair J., 2015, *Monthly Notices of the Royal Astronomical Society*, 451, 2417
- Sanidas S. A., Battye R. A., Stappers B. W., 2012, *Phys. Rev. D*, 85, 122003
- Sedoglavic A., 2002, *Journal of Symbolic Computation*, 33, 735
- Sesana A., 2013, *Classical and Quantum Gravity*, 30, 224014
- Sesana A., Vecchio A., 2010, *Phys. Rev. D*, 81, 104008
- Sesana A., Vecchio A., Colacino C. N., 2008, *Monthly Notices of the Royal Astronomical Society*, 390, 192
- Shannon R. M., Cordes J. M., 2010a, *ApJ*, 725, 1607
- Shannon R. M., Cordes J. M., 2010b, *ApJ*, 725, 1607
- Shannon R. M., et al., 2013, *The Astrophysical Journal*, 766, 5
- Simon D., 2006, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*. Wiley-Interscience, USA
- Skilling J., 2006, *Bayesian Analysis*, 1, 833
- Sornette D., 2004, *Critical phenomena in natural sciences : chaos, fractals selforganization and disorder : concepts and tools*
- Speagle J. S., 2020, *Monthly Notices of the Royal Astronomical Society*, 493, 3132
- Speri L., Porayko N. K., Falxa M., Chen S., Gair J. R., Sesana A., Taylor S. R., 2023, *MNRAS*, 518, 1802
- Spiewak R., et al., 2022, *Publ. Astron. Soc. Australia*, 39, e027
- Stairs I. H., Lyne A. G., Shemar S. L., 2000, *Nature*, 406, 484
- Stairs I. H., et al., 2019, *MNRAS*, 485, 3230
- Sykes B., Middleton H., Melatos A., Di Matteo T., DeGraf C., Bhowmick A., 2022, *MNRAS*, 511, 5241
- Tarafdar P., et al., 2022, *Publications of the Astronomical Society of Australia*, 39, e053
- Taylor S. R., Huerta E. A., Gair J. R., McWilliams S. T., 2016, *ApJ*, 817, 70
- Tiburzi C., 2018, *Publ. Astron. Soc. Australia*, 35, e013
- Trassinelli M., 2019, *Proceedings*, 33
- Urama J. O., Link B., Weisberg J. M., 2006, *MNRAS*, 370, L76
- Vargas A., Melatos A., 2023, TBD, 1, 1
- Verbiest J. P. W., Weisberg J. M., Chael A. A., Lee K. J., Lorimer D. R., 2012, *ApJ*, 755, 39
- Verbiest J. P. W., Osłowski S., Burke-Spolaor S., 2021, in , *Handbook of Gravitational Wave Astronomy*. p. 4, doi:10.1007/978-981-15-4702-7\_4-1
- Villani C., 2009, *The Wasserstein distances*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 93–111, doi:10.1007/978-3-540-71050-9\_6, [https://doi.org/10.1007/978-3-540-71050-9\\_6](https://doi.org/10.1007/978-3-540-71050-9_6)
- Wan E., Van Der Merwe R., 2000, in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373)*. pp 153–158, doi:10.1109/ASSPCC.2000.882463
- Wyithe J. S. B., Loeb A., 2003, *ApJ*, 590, 691
- Xu H., et al., 2023, *Research in Astronomy and Astrophysics*, 23, 075024
- Xue X., et al., 2021, *Phys. Rev. Lett.*, 127, 251303
- Yao J. M., Manchester R. N., Wang N., 2017, *ApJ*, 835, 29
- Yardley D. R. B., et al., 2010, *MNRAS*, 407, 669
- Zarchan P., Musoff H., 2000, *Fundamentals of Kalman Filtering: A Practical Approach*. Progress in astronautics and aeronautics, American Institute of Aeronautics and Astronautics, <https://books.google.com.au/books?id=AQxRAAAMAAJ>
- Zhu X. J., et al., 2014, *MNRAS*, 444, 3709
- Zhu X.-J., et al., 2015a, *Monthly Notices of the Royal Astronomical Society*, 449, 1650
- Zhu X.-J., et al., 2015b, *Monthly Notices of the Royal Astronomical Society*, 449, 1650
- Zhu X.-J., Wen L., Xiong J., Xu Y., Wang Y., Mohanty S. D., Hobbs G., Manchester R. N., 2016, *Monthly Notices of the Royal Astronomical Society*, 461, 1317
- Zic A., et al., 2023a, *arXiv e-prints*, p. arXiv:2306.16230
- Zic A., et al., 2023b, *arXiv e-prints*, p. arXiv:2306.16230
- van Eysden C. A., Melatos A., 2010, *MNRAS*, 409, 1253

This paper has been typeset from a  $\text{\LaTeX}$  file prepared by the author.