Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY
MNRAS **530**, 4648–4664 (2024) https://doi.org/10.1093/mnras/stae770
Advance Access publication 2024 March 26

Analysing radio pulsar timing noise with a Kalman filter: a demonstration involving PSR J1359−6038

Nicholas J. O'Neill,[1,2★] Patrick M. Meyers [3,1,2] and Andrew Melatos[1,2]
[1] *School of Physics, University of Melbourne, Parkville, VIC 3010, Australia*
[2] *Australian Research Council Centre of Excellence for Gravitational Wave Discovery (OzGrav), University of Melbourne, Parkville, VIC 3010, Australia*
[3] *Theoretical Astrophysics Group, California Institute of Technology, Pasadena, CA 91125, USA*

Accepted 2024 March 13. Received 2024 March 12; in original form 2023 December 1

ABSTRACT
In the standard two-component crust-superfluid model of a neutron star, timing noise can arise when the two components are perturbed by stochastic torques. Here it is demonstrated how to analyse fluctuations in radio pulse times of arrival with a Kalman filter to measure physical properties of the two-component model, including the crust-superfluid coupling time-scale and the variances of the crust and superfluid torques. The analysis technique, validated previously on synthetic data, is applied to observations with the *Molonglo Observatory Synthesis Telescope* of the representative pulsar PSR J1359−6038. It is shown that the two-component model is preferred to a one-component model, with log Bayes factor 6.81 ± 0.02. The coupling time-scale and the torque variances on the crust and superfluid are measured with 90 per cent confidence to be $10^{7.1^{+0.8}_{-0.5}}$ s and $10^{-24.0^{+0.4}_{-5.6}}$ rad^2 s^{-3} and $10^{-21.7^{+3.5}_{-0.9}}$ rad^2 s^{-3}, respectively.

Key words: methods: data analysis – pulsars: general – stars: neutron – stars: rotation.

1 INTRODUCTION

High-precision pulsar timing data are a rich and plentiful resource for probing the properties of neutron star interiors. They show that pulsars decelerate secularly with small, random fluctuations away from the secular trend in the form of occasional glitches (Shemar & Lyne 1996; Hobbs, Lyne & Kramer 2010; Espinoza et al. 2011; Haskell & Melatos 2015) and persistent pulsar timing noise (Shannon & Cordes 2010; Lentati et al. 2016; Parthasarathy et al. 2019; Lower et al. 2020). The random fluctuations can be related to internal physics, such as far-from-equilibrium processes involving superfluid vortices (Anderson & Itoh 1975; Warszawski & Melatos 2011), Type II superconductor flux tubes (Ruderman, Zhu & Chen 1998; Drummond & Melatos 2017), and hydrodynamic turbulence (Greenstein 1970; Melatos & Link 2014). One prominent structural theory supported by these observations is the two-component model, in which neutron stars are composed of a rigid outer crust and a superfluid core (Baym et al. 1969). The crust and superfluid are coupled by forces, which act to restore corotation and therefore damp observable fluctuations in the crust's angular velocity, as happens during post-glitch recoveries. With respect to timing noise specifically, the tendency to restore corotation through crust-superfluid coupling should be detectable statistically, for example, through the autocorrelation time-scale (Price et al. 2012). One advantage is that timing noise occurs all the time, so there is an abundance of data to work with.

The two-component model of neutron stars was originally motivated by observations of pulsar glitches (Baym et al. 1969). Glitches are rare events in which a pulsar spins up impulsively. Following the event, the spin-down rate (which sometimes also changes during the event) returns partially or wholly to the rate before the event, due to some friction-like interaction between the crust and superfluid operating over some time-scale τ (Anderson & Itoh 1975; Lamb, Pines & Shaham 1978a; Pines & Alpar 1985; Haskell & Melatos 2015; Antonelli, Montoli & Pizzochero 2022; Antonopoulou, Haskell & Espinoza 2022; Zhou et al. 2022). As timing noise also involves a perturbation of the crust's angular velocity, it is likely to either drive or be driven by glitch-like differential rotation between the crust and superfluid, which is damped by the same restoring forces that bring the two components back into corotation after a glitch (Lamb et al. 1978b; van Eysden & Melatos 2010; Price et al. 2012; Gügercinoğlu & Alpar 2017). We aim to measure the coupling between the two components by looking for statistical evidence of its associated relaxation time-scale τ in timing noise data.

Many previous timing noise studies focus on calculating ensemble averaged quantities derived from the whole data set such as the noise amplitude (Boynton et al. 1972; Groth 1975; Cordes 1980; Cordes & Helfand 1980; Shannon & Cordes 2010) and power spectrum (van Haasteren et al. 2009; Hobbs et al. 2010; Lentati et al. 2016; Parthasarathy et al. 2019; Goncharov, Zhu & Thrane 2020; Parthasarathy et al. 2020). Although ensemble averages are good for measuring some quantities, they do not make use of the information contained in the specific, time-ordered sequence of times of arrival (TOAs), that is, the specific realization of the random process presented to the observer (Vargas & Melatos 2023). In this paper, we seek to extract this extra information using a Kalman filter, a tool commonly used for signal processing in electrical engineering applications (Kalman 1960). Given a model for how a system evolves with time (here the two-component model), the Kalman filter tracks the specific random fluctuations in the observed

★ E-mail: noneill@student.unimelb.edu.au

time series and estimates the most likely state of the system at every time-step, that is, the most likely values of the dynamical variables (the crust and superfluid angular velocities) in the two-component model. Moreover, when combined with a nested sampler, the Kalman filter estimates the most likely values of the static parameters in the model, for example, the crust-superfluid coupling coefficient and the variances of the stochastic torques. Two previous papers (Meyers, Melatos & O'Neill 2021a; Meyers et al. 2021b) explained this approach and validated it with synthetic data.

In this paper, we demonstrate in principle that the approach works successfully on real data by applying it to *Molonglo Observatory Synthesis Telescope* observations of PSR J1359−6038 (Bailes et al. 2017; Lower et al. 2020). This object is chosen, because its timing noise power spectrum is similar to that predicted by the simplest version of the two-component model, as discussed below (see Section 4 and Appendix A). It also offers relatively good-quality data, with 429 TOAs over 3.5 yr of monitoring, with an average TOA uncertainty of 5.1×10^{-5}s. Other pulsars can be analysed but an analysis of a sample of objects is outside the scope of this paper, whose goal is to demonstrate the real-world applicability of the method by way of a worked example, to pave the way for fuller studies in the future.

The paper is structured as follows. In Section 2, the two-component model is introduced. In Section 3, the Kalman filter and its implementation are explained, including the question of parameter identifiability and the choice of priors based on existing observations. In Section 4, we discuss the properties of the data from PSR J1359−6038, including the number of observations and measurement uncertainties, and describe the conversion from TOAs to local frequencies. In Section 5, we present the parameter estimates for PSR J1359−6038. In Section 6, we present a Bayesian comparison between the one-component and two-component models. In Section 7, we conclude by interpreting the results and validating them with Monte Carlo posterior checks to verify their statistical significance. The appendices contain for completeness an analytic derivation of the power spectral density (PSD) of the two-component model, the explicit formulas for the matrices used in the Kalman filter, the Kalman filter recurrence relations, calculations demonstrating the accuracy of the parameter estimates as a function of data volume and measurement errors, and details of the Bayesian comparison of the two-component model with a simpler, one-component model.

2 TWO-COMPONENT MODEL

The two-component model of a neutron star consists of a rigid crust and superfluid core, which are assumed to rotate uniformly, with angular velocities Ω_c and Ω_s, respectively. The components obey the equations of motion (Baym et al. 1969; Gügercinoğlu & Alpar 2017)

$$I_c \frac{d\Omega_c}{dt} = -\frac{I_c}{\tau_c}(\Omega_c - \Omega_s) + N_c + \xi_c(t), \tag{1}$$

$$I_s \frac{d\Omega_s}{dt} = -\frac{I_s}{\tau_s}(\Omega_s - \Omega_c) + N_s + \xi_s(t), \tag{2}$$

where the subscripts 'c' and 's' label the crust and the superfluid, respectively, I_c and I_s are the moments of inertia, N_c and N_s are constant external torques, ξ_c and ξ_s are stochastic torques, and τ_c and τ_s are coupling time-scales. The astrophysical origins of the torques on the right-hand sides of equations (1) and (2) are discussed below. The stochastic torques obey Gaussian, white noise statistics with

$$\langle \xi_{c,s}(t) \rangle = 0 \tag{3}$$

$$\langle \xi_{c,s}(t)\xi_{c,s}(t') \rangle = \sigma_{c,s}^2 \delta(t - t'), \tag{4}$$

where $\langle \dots \rangle$ denotes the ensemble average, and σ_c and σ_s are noise amplitudes.

The coupling between the crust and superfluid is assumed to act like friction between the two components, reducing their relative velocity. In this paper, $|\Omega_c - \Omega_s|$ is assumed to be small enough that the coupling torque is approximately linear in this difference on the right hand sides of equations (1) and (2). Over a typical observing time-scale (decades), and over the relaxation time-scales τ_c and τ_s (typically weeks) (Price et al. 2012), the torques N_c and N_s can be approximated as constants. N_c represents the net secular external torque applied to the crust, including the electromagnetic radiation reaction torque (Goldreich & Julian 1969) and the gravitational radiation reaction torque (Ferrari & Ruffini 1969; Ostriker & Gunn 1969). N_s represents the net secular torque on the superfluid which includes electromagnetic and gravitational components, if the superfluid is threaded by a corotating magnetic field and possesses a time-varying mass or current quadrupole moment, respectively. The torque ξ_c represents the net random torque acting on the crust, which may be negligible in a non-accreting radio pulsar, except perhaps in the presence of seismic activity (Middleditch et al. 2006; Chugunov & Horowitz 2010; Giliberti & Cambiotti 2022; Kerin & Melatos 2022). ξ_s represents the net random torque acting on the superfluid, including the torque from superfluid vortex unpinning (Anderson & Itoh 1975; Warszawski & Melatos 2011). The stochastic torques drive timing noise in the observable Ω_c, even in the scenario $\xi_c = 0$ and $\xi_s \neq 0$.

Equations (1) and (2) have been generalized recently to multiple internal components by Antonelli, Basu & Haskell (2023). We restrict the analysis in this paper to two components, because the available data are insufficient to constrain a more complicated model. Indeed, the data are insufficient even to constrain equations (1) and (2) fully, as discussed in Section 3.

3 KALMAN TRACKING AND ESTIMATION

Given an observed time series $\Omega_c(t_1), \dots, \Omega_c(t_N)$, one can use a Kalman filter to infer the most probable sequence of states $\Omega_c(t_i)$ and $\Omega_s(t_i)$ $(1 \le i \le N)$ traversed by the system described by equations (1) and (2). In this paper, the Kalman filter assumes the measurements are the true state of the system plus some Gaussian measurement noise. The Kalman filter then uses the assumption that the system evolves according to equations (1) and (2) to separate out the real evolution of the system from the measurement noise on a most probable basis, which gives estimates of $\Omega_c(t_i)$ and $\Omega_s(t_i)$ at each t_i together with an error on each estimate. This procedure, termed Kalman tracking, is described in subsection 3.1. The Kalman filter's ability to recover the model parameters from Ω_c data is assessed in subsection 3.2. To estimate the parameters, one calculates the Kalman filter likelihood (of the model producing the observed data) as a function of the static parameters to find their most probable values given $\Omega_c(t_i)$. This procedure, termed Kalman estimation, is described in subsection 3.3. The results depend on the priors, which are specified in subsection 3.4 together with their astrophysical motivations.

3.1 Kalman tracking

The state of the system at time t_i is denoted by

$$\boldsymbol{X}_i = \begin{pmatrix} \Omega_c(t_i) \\ \Omega_s(t_i) \end{pmatrix}. \tag{5}$$

The Kalman filter predicts the next state of the system given the current one. Solving equations (1) and (2) gives a discrete equation for updating the state from time t_{i-1} to t_i

$$\boldsymbol{X}_i = \boldsymbol{F}_i \boldsymbol{X}_{i-1} + \boldsymbol{T}_i + \boldsymbol{w}_i. \tag{6}$$

Equation (6) multiples $\boldsymbol{X}_{i-1}$ by a transition matrix $\boldsymbol{F}_i$, which comes from the coupling terms in equations (1) and (2); adds a vector $\boldsymbol{T}_i$, which comes from the deterministic torques; and adds a random vector $\boldsymbol{w}_i$, which comes from the stochastic torques and is drawn from a Gaussian distribution with

$$\langle \boldsymbol{w}_i \rangle = \boldsymbol{0} \tag{7}$$

$$\langle \boldsymbol{w}_i \boldsymbol{w}_j \rangle = \boldsymbol{Q}_i \delta_{ij}. \tag{8}$$

Explicit expressions for $\boldsymbol{F}_i$, $\boldsymbol{T}_i$ and $\boldsymbol{Q}_i$ are given in Appendix B.

For general physical systems, the state and measurements at t_i may be related in a complicated non-linear way. In this paper, however, the relation is simple: the measured and true $\Omega_c(t_i)$ differ by an additive measurement error, and $\Omega_s(t_i)$ is hidden, that is, it cannot be measured directly at all. Mathematically, we encode this by saying that a measurement $\boldsymbol{Y}_i$ at time t_i is related to the state of the system $\boldsymbol{X}_i$ by

$$\boldsymbol{Y}_i = \boldsymbol{C}\boldsymbol{X}_i + \boldsymbol{v}_i. \tag{9}$$

In equation (9), $\boldsymbol{v}_i$ is the measurement noise and is drawn from a normal distribution with

$$\langle \boldsymbol{v}_i \rangle = \boldsymbol{0} \tag{10}$$

$$\langle \boldsymbol{v}_i \boldsymbol{v}_j \rangle = \boldsymbol{R}_i \delta_{ij}. \tag{11}$$

$\boldsymbol{C}$ is the observation matrix which determines which components of the state can be measured. In radio pulsar timing experiments, only the crust can be measured, since the core is hidden from view, implying[1]

$$\boldsymbol{C} = (1, 0). \tag{12}$$

The Kalman filter computes the expected evolution of the system through equation (6). It updates the expected evolution with new information from the measurement at t_i through equation (9), separating the process noise $\boldsymbol{w}_i$ from the measurement noise $\boldsymbol{v}_i$. The detailed implementation of this two-step, predictor–corrector algorithm is discussed in Appendix C to help the interested reader reproduce the results in Sections 5 and 6.

3.2 Identifiability

In general, there is no guarantee that all six of the static parameters τ_c, τ_s, σ_c, σ_s, N_c, N_s in equations (1) and (2) can be identified from the time series $\Omega_c(t_1), \ldots, \Omega_c(t_N)$, even if the data volume is arbitrarily large ($N \to \infty$). Certain parameters cannot be separated from the rest and can be estimated only in combination. This issue, known as identifiability, is widely recognized in electrical engineering applications. A number of formal techniques have been developed to handle it, as summarized by Bellman & Åström (1970) for example. In this section, we analyse what combinations of the static parameters in equations (1) and (2) are identifiable, in order to interpret the posterior distribution.

[1]In the future, gravitational waves radiated from the core may be detectable, whereupon $\Omega_s(t_i)$ could be measured directly.

As a starting point, it is instructive to express equations (1) and (2) as an equation for Ω_c on its own, viz.

$$\ddot{\Omega}_c = -\left(\frac{1}{\tau_c} + \frac{1}{\tau_s}\right)\dot{\Omega}_c + \left(\frac{N_c}{\tau_s I_c} + \frac{N_s}{\tau_c I_s}\right) + \frac{\xi_c}{\tau_s I_c} + \frac{\xi_s}{\tau_c I_s} + \frac{\dot{\xi}_c}{I_c}. \tag{13}$$

Equation (13) determines $\Omega_c(t)$ fully, supplemented by initial conditions $\Omega_c(t_1)$ and $\dot{\Omega}_c(t_1)$. Therefore the static parameter combinations that appear in equation (13) are identifiable – that is, they can be estimated from the data – but they cannot be decomposed into their elements. For example, τ_c and τ_s cannot be estimated separately, because they appear in the irreducible combination $\tau_c^{-1} + \tau_s^{-1}$ in the first term on the right-hand side of equation (13), and no additional, independent equation of motion exists beyond equation (13), in which τ_c and τ_s appear separately.

The first term on the right-hand side of equation (13) contains the parameter

$$\tau = \frac{\tau_c \tau_s}{\tau_c + \tau_s}, \tag{14}$$

which is a combined relaxation time-scale for the system as a whole. The second term (when multiplied by τ) contains the parameter

$$\langle \dot{\Omega}_c \rangle = \frac{1}{\tau_c + \tau_s}\left(\tau_c \frac{N_c}{I_c} + \tau_s \frac{N_s}{I_s}\right), \tag{15}$$

which is the ensemble-averaged frequency derivative of the pulsar. The noiseless evolution of Ω_c is governed solely by τ and $\langle \dot{\Omega}_c \rangle$ so these parameters are easiest to estimate. For notational convenience, and to assist with physical interpretation, we also introduce the complementary parameter combinations

$$r = \frac{\tau_s}{\tau_c} \tag{16}$$

$$\langle \Omega_c - \Omega_s \rangle = \tau \left(\frac{N_c}{I_c} - \frac{N_s}{I_s}\right). \tag{17}$$

In equations (16) and (17), r is the ratio of relaxation time-scales, which equals I_s/I_c when the coupling torques form an action-reaction pair, and $\langle \Omega_c - \Omega_s \rangle$ is the ensemble-averaged angular velocity lag between the crust and superfluid.

It is less straightforward to read off by sight whether the noise amplitudes $Q_c = \sigma_c^2/I_c^2$ and $Q_s = \sigma_s^2/I_s^2$ can be estimated by the Kalman filter. However, in Appendix D, the question is answered empirically using synthetic data. It turns out that the Kalman filter can estimate Q_c reliably (in addition to τ and $\langle \dot{\Omega}_c \rangle$) but usually not Q_s. An analytic treatment of the identifiability of Q_c and Q_s is given in Appendix E. The analysis suggests that it is easiest to recover the parameter combinations r, τ, Q_c, Q_s, $\langle \Omega_c - \Omega_s \rangle$, $\langle \dot{\Omega}_c \rangle$, and we switch to them in what follows.

3.3 Kalman filter likelihood

We construct a likelihood function $p(\boldsymbol{Y}_{1:N}|\boldsymbol{\theta})$ with $\boldsymbol{Y}_{1:N} = \{\boldsymbol{Y}_1, \boldsymbol{Y}_2, \ldots, \boldsymbol{Y}_N\}$ for the data given a choice of parameters. We calculate $p(\boldsymbol{Y}_{1:N}|\boldsymbol{\theta})$ from the optimal state sequence output by the Kalman filter. Intuitively, if the parameters are chosen well (i.e. near their true values), the model's trajectory in time matches the data closely, and $p(\boldsymbol{Y}_{1:N}|\boldsymbol{\theta})$ is relatively large.

Let $\boldsymbol{\epsilon}_i$ be the difference between the measurement and the prediction of the state at t_i, and let $\boldsymbol{S}_i$ be the covariance of $\boldsymbol{\epsilon}_i$. Then

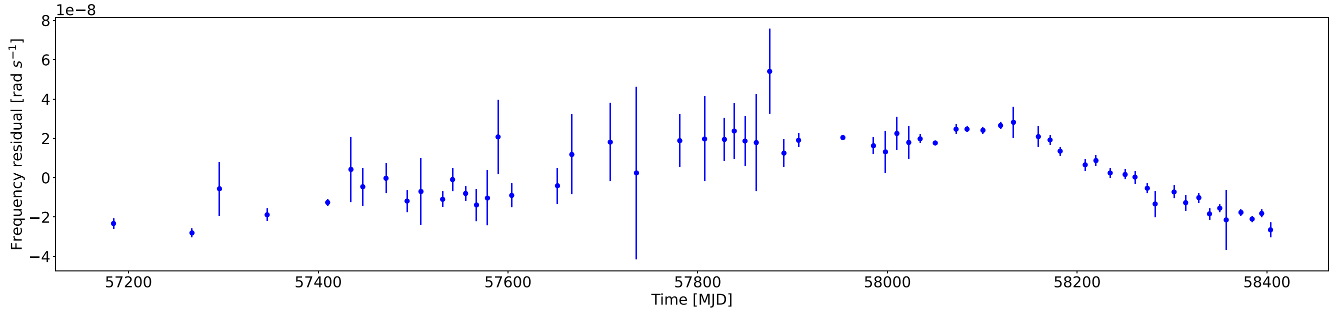

Figure 1. Residuals of frequencies fitted to 428 TOAs versus time (units: MJD) for the PSR J1359−6038 after subtracting a linear trend ($\Omega(t)$[rad s^{-1}] = 49.2888 − 2.4472 × $10^{-12}t$[s]) and removing two outliers due to poor fitting. The vertical scale is in units of 10^{-8} rad s^{-1}. The frequencies and error bars are calculated using TEMPO2 from the TOAs and their uncertainties in the UTMOST data release.

Bayes's rule gives the posterior distribution of $\boldsymbol{\theta}$ as

$$\ln p(\boldsymbol{\theta}|\boldsymbol{Y}_{1:N}) = -\frac{1}{2}\sum_{i=1}^{N-1}\left[N_Y \ln(2\pi) + \ln\det(\boldsymbol{S}_i) + \boldsymbol{\epsilon}_i^T \boldsymbol{S}_i^{-1}\boldsymbol{\epsilon}_i\right] + \ln p(\boldsymbol{\theta}) - \ln(Z), \quad (18)$$

where N_Y is the dimension of $\boldsymbol{Y}_i$, which in our case is always unity, $p(\boldsymbol{\theta})$ is the prior distribution of $\boldsymbol{\theta}$, and Z is the evidence, which acts as a normalization constant. Appendix C supplies more details on the derivation of equation (18). The `dynesty` sampler (Speagle 2020) is used to efficiently sample the posterior distribution.

3.4 Prior distribution

In the absence of other information, we choose log-uniform priors for the parameters that are positive definite and plausibly span several decades, namely r, τ, Q_c, and Q_s. We choose uniform priors for the parameters with small ranges, namely $\langle\dot{\Omega}_c\rangle$ and $\langle\Omega_c - \Omega_s\rangle$. As more pulsars are analysed in the future, and a picture of the distribution of (say) τ across the population emerges, it may become appropriate to use more informative priors in future work.

We can use previous measurements of glitch relaxation time-scales to infer reasonable bounds on τ, motivated by equation (14). The data in Yu et al. (2013) imply $10^5 \le \tau/(1\ \mathrm{s}) \le 10^8$. Models of stellar structure suggest $10^{-2} \le r = \tau_s/\tau_c = I_s/I_c \le 10^2$ (Link, Epstein & Lattimer 1999; Lyne, Shemar & Smith 2000; Espinoza et al. 2011; Chamel 2012). Reasonable bounds on Q_c and Q_s follow from previous, independent measurements of timing noise amplitude. For the crust one measures typically $10^{-30} \le Q_c/(1\ \mathrm{rad}^2\ \mathrm{s}^{-3}) \le 10^{-16}$ (Cordes & Downs 1985; Çerri-Serim et al. 2019; Lower et al. 2020; Vargas & Melatos 2023). In the absence of additional information, we assume the same range for Q_s, noting that radio pulsar timing experiments cannot track the core to measure Q_s directly. For $\langle\dot{\Omega}_c\rangle$, a central estimate $\langle\dot{\Omega}_c\rangle_{\mathrm{cent}}$ and an uncertainty σ can be obtained by fitting a linear trend to the time series $\Omega_c(t_i)$. We adopt a uniform prior for $\langle\dot{\Omega}_c\rangle$ with a range of $[\langle\dot{\Omega}_c\rangle_{\mathrm{cent}} - 1000\sigma, \langle\dot{\Omega}_c\rangle_{\mathrm{cent}} + 1000\sigma]$. For the crust–core lag, one extreme case $N_c/I_c \ll N_s/I_s \le 0$ implies $\langle\Omega_c - \Omega_s\rangle \sim \tau\langle\dot{\Omega}_c\rangle$. The opposite extreme $N_s/I_s \ll N_c/I_c \le 0$ implies $\langle\Omega_c - \Omega_s\rangle \sim -\tau\langle\dot{\Omega}_c\rangle$. We therefore assume $\tau_{\max}\langle\dot{\Omega}_c\rangle_{\min} \le \langle\Omega_c - \Omega_s\rangle \le -\tau_{\max}\langle\dot{\Omega}_c\rangle_{\min}$, where the subscripts 'max' and 'min' denote the maximum and minimum values of the prior range given earlier in this section. For most pulsars, one finds $-10^{-11} \le \langle\dot{\Omega}_c\rangle \le 0$ (Helfand et al. 1980) and hence $-10^{-3} \le \langle\Omega_c - \Omega_s\rangle \le 10^{-3}$. For more details on justifying the prior distribution, the reader is referred to Meyers et al. (2021a); Meyers et al. (2021b).

4 DATA

The data used in this paper come from the UTMOST pulsar observing programme[2] (Lower et al. 2020). They are in the form of barycentred pulse TOAs. To assist parameter estimation, we select a test object with a relatively large number of TOAs with relatively small error bars. We also avoid objects exhibiting glitches, because glitches are not included in the dynamical model equations (1) and (2). An object which satisfies these criteria is PSR J1359−6038. The UTMOST data release for PSR J1359−6038 contains 429 TOAs, one of which we discard as an outlier.[3] The average TOA error is 51 μs. The TOAs span 1263 d. In Lower et al. (2020) the timing noise in PSR J1359−6038 is measured to have a spectral index of $\gamma = -5.1^{+0.8}_{-0.9}$, that is, the PSD of the phase residuals scales $\propto f^\gamma$ at $f \gtrsim 10^{-8}$ Hz. This is comparable to the theoretical prediction for the two-component model equations (1) and (2), of $\gamma = -4$, given in Appendix equation (A13).

The Kalman filter in Section 3 ingests pulse frequency data. Hence, the phase information in the TOAs must be converted to local frequencies $\Omega_c(t_1), \ldots, \Omega_c(t_N)$. This is done using the standard pulsar timing software package, TEMPO2 (Edwards, Hobbs & Manchester 2006; Hobbs, Edwards & Manchester 2006). To generate a local frequency estimate, we feed a set, S_n, of consecutive TOAs into TEMPO2 and extract $\Omega(\bar{t}_n)$ and its associated error, where $\bar{t}_n$ is the average of the TOAs in the set S_n. We construct the disjoint sets, $S_1, S_2, \ldots, S_N$, from the TOAs, $t_1, t_2, \ldots, t_M$ ($M \ge N$), starting with t_1, according to the following two rules: (i) S_n contains t_i and all $t_j > t_i$ with $|t_j - t_i| < 10$ days and (ii) there must be at least three TOAs per set, and the 10-day window is lengthened when necessary to achieve this.

The fitting process yields 62 frequency data points spanning 1220 d (the $\bar{t}_n$ values span a slightly shorter time than the t_i values). We find $\Omega_c(t_i) = \Omega_c(t_1) + (t_i - t_1)\dot{\Omega}_c(t_1)$ to an excellent approximation, with $\Omega_c(t_1) = 49.2767$ rad s^{-1} and $\dot{\Omega}_c(t_1) = -2.4472 \times 10^{-12}$ rad s^{-2}. Subtracting the linear, long-term trend yields the frequency residuals plotted in Fig. 1. The standard deviation of the residuals is 1.7×10^{-8} rad s^{-1}. The mean uncertainty is 7.8×10^{-9} rad s^{-1}. However the size of the error bars varies with the errors of the original TOAs, the number of TOAs used in a fit, and the time span of a fit. The interval from MJD 57 400 to MJD 57 900 features TOAs

[2] https://github.com/Molonglo/TimingDataRelease1/

[3] The TOA at MJD 58190.7 differs from the trend of its neighbouring TOAs by ≈ 8 times more than its error bars. An investigation in Dunn et al. (2022) suggests that the outlier was due to observatory conditions since a similar offset occurs in data for other pulsars at the same observatory at that time.

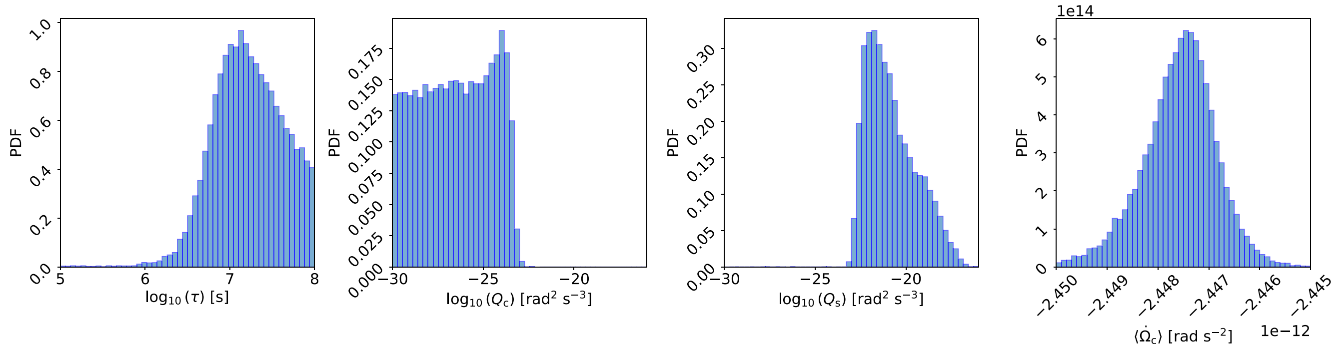

Figure 2. Marginalized posterior distributions for the four parameters, τ, $Q_c = \sigma_c^2/I_c^2$, $Q_s = \sigma_s^2/I_s^2$ and $\langle\dot{\Omega}_c\rangle$ that can be estimated reliably for the two-component model applied to PSR J1359−6038. The marginalized posteriors are plotted as histograms spanning the prior range, except for $\langle\dot{\Omega}_c\rangle$ where only the region around the peak is shown. The units of the parameters are specified on the horizontal axis of each panel. The vertical axis carries a linear scale, showing the value of the marginalized PDF for that parameter.

with relatively large error bars and spacings; the median uncertainty on $\Omega(t_i)$ therein is 1.3×10^{-8} rad s^{-1} compared to 2.6×10^{-9} rad s^{-1} for the measurements outside that interval. Some fitted frequencies deviate significantly from the linear trend, possibly due to an issue with fitting to TOAs with large errors. We remove two outliers with residuals from the trend below -10^{-7} rad s^{-1}.

The conversion from TOAs to frequencies is imperfect. Fitting frequencies filters the timing noise on the time-scale over which the fit is done, so some information is lost. The effect of filtering is tested on simulated TOAs in Appendix F. We find that it causes Q_c to be slightly underestimated and makes τ harder to recover.

5 ESTIMATED PARAMETERS OF THE TWO-COMPONENT MODEL

In this section, we apply the Kalman tracking and estimation procedure described in Section 3 to the UTMOST data described in Section 4. Marginalized posterior distributions for the static parameters that can be identified reliably, namely τ, Q_c, Q_s, and $\langle\dot{\Omega}_c\rangle$, are presented in subsection 5.1. The four parameters are discussed individually in greater detail in subsections 5.2–5.4 and interpreted astrophysically. The two parameters that cannot be identified reliably, namely r and $\langle\Omega_c - \Omega_s\rangle$, are discussed in subsection 5.5.

5.1 Joint posterior distribution

The six-dimensional joint posterior distribution calculated from the PSR J1359−6038 frequency data in Fig. 1 using equation (18) is displayed in the traditional format of a corner plot in Appendix G. Four parameters can be estimated reliably and exhibit well-formed peaks: τ, Q_c, Q_s, and $\langle\dot{\Omega}_c\rangle$. Two parameters, r and $\langle\Omega_c - \Omega_s\rangle$, cannot be estimated reliably and rail against the prior bounds.

One-dimensional cross-sections of the posterior, marginalized over all but one parameter, are displayed as histograms in Fig. 2 for the four parameters that can be estimated reliably. The marginalized posteriors in Fig. 2 are unimodal, with the peaks falling comfortably within the respective prior ranges in Table 1. There are distinct peaks for τ, Q_c, Q_s, and $\langle\dot{\Omega}_c\rangle$. They occur at $\tau = 1.3 \times 10^7$ s, $Q_c = 1.05 \times 10^{-24}$ rad^2 s^{-3}, $Q_s = 1.82 \times 10^{-22}$ rad^2 s^{-3} and $\langle\dot{\Omega}_c\rangle = -2.4475 \times 10^{-12}$ rad s^{-2}.

The widths of the 90 per cent confidence intervals of the histograms in Fig. 2 are 1.3 dex, 6.0 dex, 4.4 dex, and 2.4×10^{-15} rad s^{-2} for τ, Q_c, Q_s, $\langle\dot{\Omega}_c\rangle$, respectively. The 90 per cent confidence intervals cover 43 per cent, 43 per cent, 31 per cent, and 1.4 per cent of their respective prior ranges, confirming that the four parameters are identified reliably.

Table 1. Prior distributions introduced in subsection 3.4 for the parameter estimation carried out in Appendix Figs G1 and H1. The rightmost column indicates the prior distribution, $p(\cdot)$, on each parameter. $\mathcal{U}(a, b)$ indicates a uniform distribution between a and b. The priors of the first four (last two) parameters are log-uniform (uniform).

Parameter	Units	Prior
r		$\log\mathcal{U}(10^{-2}, 10^2)$
τ	s	$\log\mathcal{U}(10^5, 10^8)$
Q_c	rad^2 s^{-3}	$\log\mathcal{U}(10^{-30}, 10^{-16})$
Q_s	rad^2 s^{-3}	$\log\mathcal{U}(10^{-30}, 10^{-16})$
$\langle\Omega_c - \Omega_s\rangle$	rad s^{-1}	$\mathcal{U}(-10^{-3}, 10^{-3})$
$\langle\dot{\Omega}_c\rangle$	rad s^{-2}	$\mathcal{U}(-2.5323 \times 10^{-12}, -2.3621 \times 10^{-12})$

Numerical values for the parameter estimates for PSR J1359−6038 conditional on the two-component model given by equations (1) and (2) are summarized for completeness in Table 2.

5.2 Coupling time-scale

In Fig. 2, we find that the marginalized posterior for τ has a well-defined peak. With 90 per cent confidence we recover $3.8 \times 10^6 \leq \tau/(1\,\mathrm{s}) \leq 7.6 \times 10^7$. This result is broadly consistent with previous measurements of timing noise and post-glitch relaxation time-scales. Price et al. (2012) computed the autocorrelation time-scale of timing residuals in the objects PSR B1133+16 and PSR B1933+16 and obtained ≈ 10 days and ≈ 20 days, respectively. Post-glitch relaxation time-scales have been measured previously by many authors (McCulloch et al. 1990; Alpar et al. 1993; Shemar & Lyne 1996; Lyne et al. 2000; Wang et al. 2000; Wong, Backer & Lyne 2001; Dodson, McCulloch & Lewis 2002; Yuan et al. 2010; van Eysden & Melatos 2010; Espinoza et al. 2011; Yu et al. 2013; Espinoza et al. 2021; Lower et al. 2021; Gügercinoğlu et al. 2022). They mostly range from ≈ 10 days to ≈ 3×10^2 days and occasionally reach as high as ≈ 10^3 days. The 90 per cent confidence interval in Fig. 2 overlaps this range. The prior on τ in Table 1 is motivated by glitch observations, so the overlap is expected, but it is significant that the posterior peaks comfortably within the prior range.

Table 2. Static parameters inferred by the Kalman tracker and nested sampler for the two-component model, extracted from the six-dimensional joint posterior in Appendix Fig. G1. The τ, Q_c, Q_s, and $\langle\dot{\Omega}_c\rangle$ rows correspond to the four identifiable parameters, whose marginalized posteriors are plotted in Fig. 2. The r and $\langle\Omega_c - \Omega_s\rangle$ rows correspond to the unidentifiable parameters discussed in subsection 5.5. The fourth and fifth columns list two complementary measures of the dispersion in the posterior, viz. the full-width half-maximum (FWHM) and 90 per cent confidence intervals, respectively.

Parameter	Units	Peak value	FWHM interval	90 per cent confidence interval
$\log_{10}r$		-0.6	$(-2.0, 2.0)$	$(-1.8, 1.8)$
$\log_{10}\tau$	s	7.1	$(6.7, 7.9)$	$(6.6, 7.9)$
$\log_{10}Q_c$	rad^2 s^{-3}	-24.0	$(-30.0, -23.3)$	$(-29.6, -23.6)$
$\log_{10}Q_s$	rad^2 s^{-3}	-21.7	$(-22.7, -19.9)$	$(-22.6, -18.2)$
$\langle\Omega_c - \Omega_s\rangle$	rad s^{-1}	3×10^{-4}	$(-1.0 \times 10^{-3}, 1.0 \times 10^{-3})$	$(-9.0 \times 10^{-4}, 9.0 \times 10^{-4})$
$\langle\dot{\Omega}_c\rangle$	rad s^{-2}	-2.4475×10^{-12}	$(-2.4482 \times 10^{-12}, -2.4468 \times 10^{-12})$	$(-2.4489 \times 10^{-12}, -2.4465 \times 10^{-12})$

5.3 Torque noise amplitudes

The marginalized posteriors for Q_c and Q_s are both unimodal. With 90 per cent confidence we recover $\log_{10}(Q_c/1\,\mathrm{rad}^2\,\mathrm{s}^{-3}) = -24.0^{+0.4}_{-5.6}$ and $\log_{10}(Q_s/1\,\mathrm{rad}^2\,\mathrm{s}^{-3}) = -21.7^{+3.5}_{-0.9}$.

In previous timing noise analyses, the statistics of the timing residuals are modelled by a PSD of the form (Lentati et al. 2016; Parthasarathy et al. 2019; Lower et al. 2020)

$$P(f) = \frac{A^2}{12\pi^2 f_{\mathrm{yr}}^3}\left(\frac{f}{f_{\mathrm{yr}}}\right)^{\gamma}, \qquad (19)$$

where A^2 is a dimensionless squared amplitude, γ is a dimensionless exponent, and we define the reference frequency $f_{\mathrm{yr}} = (1\,\mathrm{yr})^{-1}$.

For $f \gg \tau^{-1} \sim 10^{-7}$ Hz, the two-component model given by equations (1)–(4) is consistent with equation (19), as demonstrated analytically in Appendix A, with $\gamma = -4$ and $\log_{10}A = 0.5(\log_{10}Q_c + 2.99)$ [see Appendix equation (A15)]. Hence the 90 per cent confidence interval $-29.6 \leq \log_{10}(Q_c/1\,\mathrm{rad}^2\,\mathrm{s}^{-3}) \leq -23.6$ in the second panel of Fig. 2 converts into $-13.3 \leq \log_{10}A \leq -10.3$. This is marginally below the 95 per cent confidence interval measured independently by Lower et al. (2020) for PSR J1359−6038, viz. $\log_{10} A = -10.0^{+0.2}_{-0.1}$, noting a slight discrepancy at the 95 per cent-confidence level in the exponent, viz. $\gamma = -5.1^{+0.8}_{-0.9}$ versus $\gamma = -4$. More generally, the estimate in Fig. 2 is comparable broadly with population studies of ordinary and millisecond pulsars, which yield $-15 \leq \log_{10}A \leq -4.9$ and $-20 \leq \gamma \leq -0.40$ (Parthasarathy et al. 2019; Lower et al. 2020; Keith & Niţu 2023). It is also comparable to but higher than population studies of millisecond pulsars only, which yield $-17 \leq \log_{10}A \leq -12$ and $-7.5 \leq \gamma \leq -0.44$ (Lentati et al. 2016; Goncharov et al. 2020; Goncharov et al. 2021). The latter result is expected, as ordinary pulsars such as PSR J1359−6038 are known to be noisier typically than millisecond pulsars. Other statistics used in the literature to quantify timing noise strength such as $\sigma_z(\tau)$ (Matsakis, Taylor & Eubanks 1997; Hobbs et al. 2010), $\sigma_{\mathcal{R}}(m, T)$ (Cordes & Helfand 1980) and $\Delta_8(t)$ (Arzoumanian et al. 1994; Shannon & Cordes 2010), are harder to compare to Q_c and are not compared here.

The detailed shapes of the peaks in Q_c and Q_s and to a lesser extent their positions are influenced by the TEMPO2 fitting process which constructs local frequencies from TOAs, as described in Section 4. To calibrate for this effect, we conduct Monte Carlo simulations with synthetic data comparing the Q_c and Q_s estimates inferred from local TEMPO2 TOA fits versus direct frequency data. The results are presented in Appendix F. When calculating local frequencies from TOAs, TEMPO2 fits a linear frequency model $\propto \dot{\Omega}_c(t_n - t_{n-1})$, which does not make any allowance for a random walk in the interval $t_{n-1} \leq t \leq t_n$. Consequently TEMPO2 acts as a low-pass filter, which smooths out features on inter-TOA time-scales. If TOAs are spaced widely enough, so that one has $\tau \ll t_n - t_{n-1}$, then TEMPO2 smooths out some evidence for τ as well. Smoothing may also bias Q_c and Q_s (Deeter 1984; Meyers et al. 2021a). Because equations (1) and (2) are linear differential equations, the timing noise in Ω_c can be written as a linear combination of the timing noise from ξ_c and ξ_s. These two contributions are separated and plotted in Fig. 3 for simulated data. The contribution from ξ_c is rougher than from ξ_s. This is because ξ_c appears directly in equation (1) for $d\Omega_c/dt$, whereas ξ_s only affects Ω_c indirectly through its time integral in Ω_s in the coupling term. Hence smoothing the data nullifies ξ_c fluctuations more than ξ_s fluctuations, overestimating Q_s and underestimating Q_c.

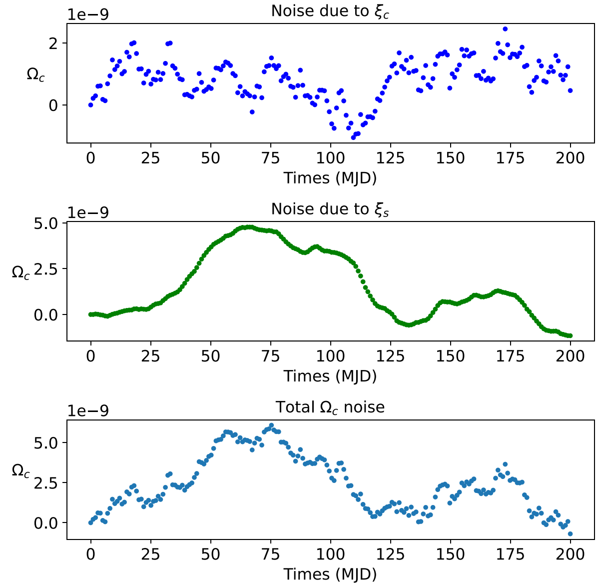

Figure 3. Residuals of simulated pulsar frequencies to illustrate the qualitative difference between the timing noise produced by ξ_c and ξ_s. The first panel shows the fluctuations in Ω_c for a simulated pulsar with $\xi_c \neq 0$ but $\xi_s = 0$, the second panel shows Ω_c produced for $\xi_s \neq 0$ and $\xi_c = 0$ and the third panel is for a simulated pulsar, where ξ_c and ξ_s are calculated by adding the fluctuations from the first and second panels.

5.4 Average spin-down rate

$\langle\dot{\Omega}_c\rangle$ is clearly recoverable with a narrow peak. The right-most panel of Fig. 2 yields $-2.4489 \times 10^{-12} \leq \langle\dot{\Omega}_c\rangle/(1\,\mathrm{rad\,s}^{-2}) \leq -2.4465 \times 10^{-12}$ for the 90 per cent confidence interval, which is consistent with the value of $\dot{\Omega}_c(t_1) = -2.4472 \times 10^{-12}$ rad s^{-2} predicted by Lower et al. (2020). The narrow peak is expected; $\langle\dot{\Omega}_c\rangle$ can be obtained by a linear regression in TEMPO2 without resorting to a Kalman filter, as demonstrated by decades of pulsar timing experiments.

5.5 Unidentifable parameters

The marginalized posteriors for r and $\langle\Omega_c - \Omega_s\rangle$ are not sharply peaked, as is apparent from columns 1 and 5 in the corner plot in Appendix Fig. G1 in Appendix G. There is no evidence of railing against the prior bounds, but the marginalized posterior is flat and therefore uninformative; the probability at the extremes of the prior range is ≈ 80 per cent and ≈ 90 per cent of the peak probability for r and $\langle\Omega_c - \Omega_s\rangle$, respectively.

The difficulty in recovering r and $\langle\Omega_c - \Omega_s\rangle$ is predicted by the identifiability analysis in subsection 3.2. The deterministic part of equation (13) does not feature r and $\langle\Omega_c - \Omega_s\rangle$. Relatedly, the simulations in Appendix D show the recovery of the six model parameters on simulated Ω_c data and confirm that r and $\langle\Omega_c - \Omega_s\rangle$ are poorly recovered. The average distances of the recovered $\log_{10}r$ and $\langle\Omega_c - \Omega_s\rangle$ parameters in Appendix Fig. D1 from their true values as a fraction of their prior widths are 41 per cent and 32 per cent, respectively.

6 ONE-COMPONENT MODEL

One may ask whether the two-component model described by equations (1) and (2) is needlessly elaborate, despite its sound phenomenological motivation. Can a simpler stochastic model explain the spin wandering statistics measured in PSR J1359−6038? The question becomes especially pertinent, when one acknowledges the difficulty in identifying r, Q_s, and $\langle\Omega_c - \Omega_s\rangle$, as seen in Appendix Figs D1 and G1 in Appendices D1 and G. It is possible, at least in principle, that the challenges of identification arise because the model is more complicated than it needs to be.

To test the above hypothesis, we repeat the Kalman filter analysis in Sections 3–5 for a one-component model. The single component corresponds to the crust (subscript 'c'), which is phase-locked to the radio pulsations. The one-component equation of motion takes the form $I_c d\Omega_c/dt = N_c + \xi_c(t)$, analogous to equation (1) but without the crust-superfluid coupling, where $\xi_c(t)$ is a Langevin driving torque; see Appendix H for details. Parameter estimates for the model parameters, namely Q_c and $\langle\dot{\Omega}_c\rangle$, are presented in the format of a traditional corner plot in Appendix H; see Appendix Fig. H1 and Table H1. The posterior peaks at $\log_{10} Q_c/(1\ \mathrm{rad}^2\mathrm{s}^{-3}) = -22.73^{+0.23}_{-0.19}$ and $\langle\dot{\Omega}_c\rangle = -2.4473 \times 10^{-12} \pm 7 \times 10^{-16}$ rad s^{-2} (90 per cent confidence intervals).

The $\langle\dot{\Omega}_c\rangle$ value recovered above is similar to that recovered for the two-component model, because it is insensitive to how the timing noise is modelled. However, the recovered Q_c for the one-component model is larger than for the two-component model by a factor of $10^{1.3} \approx 20$. This is likely because Q_c and Q_s combine through the crust-superfluid coupling to generate the noise in $\Omega_c(t)$ in the two-component model, whereas only Q_c is responsible in the one-component model, and we find $Q_s \sim 10^2 Q_c$ in Fig. 2 and Table 2 for the two-component model.

We can compare Q_c for the one-component model to the PSD normalization A measured by other researchers as in subsection 5.3. The one-component estimate of Q_c in Appendix Table H1 converts to $\log_{10} A = -9.87^{+0.12}_{-0.09}$. This is consistent with the estimate for PSR J1359−6038 in Lower et al. (2020) and with estimates for ordinary and millisecond pulsars more broadly (Parthasarathy et al. 2019; Lower et al. 2020; Keith & Niţu 2023).

A Bayesian model comparison shows that the two-component model is strongly preferred in a Bayesian sense, with a $\log_{10}$ Bayes factor of 6.81 ± 0.02 relative to the one-component model. Further details can be found in Appendix H.

7 CONCLUSION

Traditionally, timing noise studies proceed by comparing the measured variance in the phase residuals to that predicted by a microphysical or phenomenological model (Boynton et al. 1972; Groth 1975; Cordes 1980; Cordes & Helfand 1980; Shannon & Cordes 2010; Melatos & Link 2014), fitting a microphysical or phenomenological model to the PSD of the phase residuals (van Haasteren et al. 2009; Hobbs et al. 2010; Lentati et al. 2016; Parthasarathy et al. 2019; Goncharov et al. 2020; Parthasarathy et al. 2020), or measuring a relaxation time-scale using the autocorrelation function of the phase residuals (Price et al. 2012). These approaches raise interesting questions about the type of variance measured, for example Allan variance (Matsakis et al. 1997; Hobbs et al. 2010; Shannon & Cordes 2010; Melatos & Link 2014), or biases in constructing the PSD (Coles et al. 2011; van Haasteren & Levin 2013; Keith & Niţu 2023). In this paper we fit a model directly to the time-ordered $\Omega_c(t)$ data without averaging implicitly over an ensemble, that is, without analysing the phase residual PSD. The extra information made available by the analysis of a unique, time-ordered, random realization makes it feasible to estimate reliably four of the six static parameters in the classic, two-component, crust-superfluid model of a neutron star, and to distinguish statistically between one- and two-component models, with interesting astrophysical implications.

In this paper, the Kalman filter parameter estimation method developed originally by Meyers et al. (2021a); Meyers et al. (2021b) is applied to real astronomical data from a single, accurately timed object, namely PSR J1359−6038, with the aim of demonstrating the practical effectiveness of the method. The posterior distribution for the six two-component model parameters is shown in Appendix Fig. G1. Four of the six parameters are recovered reliably from the data. The results are summarized in Fig. 2 and Table 2. The peak estimates are $\tau = 1.3 \times 10^7$ s, $Q_c = 1.05 \times 10^{-24}$ rad^2 s^{-3}, $Q_s = 1.82 \times 10^{-22}$ rad^2 s^{-3} and $\langle\dot{\Omega}_c\rangle = -2.4475 \times 10^{-12}$ rad s^{-2}. The associated 90 per cent confidence intervals are $6.6 \le \log_{10}\tau/(1\ \mathrm{s}) \le 7.9$, $-29.6 \le \log_{10} Q_c/(1\ \mathrm{rad}^2\ \mathrm{s}^{-3}) \le -23.6$, $-22.6 \le \log_{10} Q_s/(1\ \mathrm{rad}^2\ \mathrm{s}^{-3}) \le -18.2$ and -2.4489×10^{-12} rad s^{-2} $\le \langle\dot{\Omega}_c\rangle \le -2.4465 \times 10^{-12}$ rad s^{-2}. The inferred coupling time-scale is broadly consistent with independent measurements based on autocorrelating phase residuals (Price et al. 2012) or fitting exponential post-glitch recoveries (Gügercinoğlu et al. 2022). The torque noise amplitudes are broadly consistent with independent measurements based on the phase residuals PSD (Parthasarathy et al. 2019; Lower et al. 2020; Keith & Niţu 2023). Two of the six two-component model parameters, namely r and $\langle\Omega_c - \Omega_s\rangle$, cannot be measured reliably, in line with the formal identifiability analysis in subsection 3.2 and Appendix E. The estimates of τ, Q_c, and Q_s, once extended to more pulsars, are likely to help illuminate the physical origin of timing noise and the nature of the spin-down-driven, far-from-equilibrium mechanisms which drive stochasticity in neutron star interiors, such as starquakes (Middleditch et al. 2006; Chugunov & Horowitz 2010; Giliberti & Cambiotti 2022; Kerin & Melatos 2022), superfluid instabilities (Andersson, Comer & Prix 2003; Melatos & Link 2014), magnetospheric fluctuations (Cheng 1987), and superfluid vortex avalanches (Anderson & Itoh 1975; Warszawski & Melatos 2011).

Bayesian model selection shows that the classic two-component model in Section 2 fits the data better than the representative one-component model given by Appendix equations (H1)–(H3), with a $\log_{10}$ Bayes factor of 6.81 ± 0.02. The Kalman filter's ability to recover successfully τ, Q_c, and Q_s for the two-component model, noting that τ and Q_s do not feature in the one-component model, adds support to the conclusion of the model selection exercise.

The results in this paper exemplify the usefulness in astrophysics of parameter estimation methods based on Kalman filtering and similar algorithms (Meyers et al. 2021a). In the future, more data and more sophisticated models will give more accurate parameter estimates. The Kalman filter can be rewritten easily to track using physical models other than equations (1) and (2). A non-linear torque can be incorporated to calculate the braking index, if the data volume is sufficient (Vargas & Melatos 2023). The model currently assumes that ξ_c and ξ_s are white noise torques, which implies that the PSD of Ω_c fluctuations scales as f^{-4} at large f, a property which is satisfied approximately by some but not all pulsars (Lower et al. 2020; Antonelli et al. 2023; Keith & Niţu 2023). Generalizing ξ_c and ξ_s to coloured noise is a standard procedure in electrical engineering (Gelb 1974). Bayesian model selection between these physically motivated alternatives is straightforward too, because the Kalman tracker and nested sampler in this paper together generate the Bayesian evidence as a by-product of the analysis.

The further step in this investigation is to extend the Kalman tracker so that it operates on TOAs directly instead of converting them first to local frequencies $\Omega_c(t_i)$ using TEMPO2. Monte Carlo simulations in Appendix F show that local TEMPO2 computation of $\Omega_c(t_i)$ biases the static parameter estimation results, underestimating Q_c and making τ harder to infer. This is not a fault with TEMPO2; it is a general consequence of the low-pass filtering introduced by any local frequency fitting process. Once the Kalman tracker is extended to operate on TOAs directly, it will be appropriate to apply the method to a wider selection of pulsars beyond PSR J1359−6038 and conduct population studies of the recoverable quantities τ, Q_c, and Q_s, which are important physically and are measured in only a few objects to date. We note in closing that the Kalman tracker and nested sampler in this paper are easy to implement and quick to run. By way of calibration, the PSR J1359−6038 analysis in this paper takes of order 1 h to run on $\sim 10^3$ TOAs. More generally, the run time scales in direct proportion to the number of TOAs.

ACKNOWLEDGEMENTS

The authors thank Liam Dunn for discussions on using the TEMPO2 software to analyse pulsar timing data. Parts of this research were conducted by the Australian Research Council Centre of Excellence for Gravitational Wave Discovery (OzGrav), through project number CE170100004. NJO is the recipient of a Melbourne Research Scholarship. The numerical calculations were performed on the OzSTAR supercomputer facility at Swinburne University of Technology. The OzSTAR programme receives funding in part from the Astronomy National Collaborative Research Infrastructure Strategy (NCRIS) allocation provided by the Australian Government.

8 DATA AVAILABILITY

The pulsar timing data for PSR J1359−6038 come from Lower et al. (2020) and are available at https://github.com/Molonglo/TimingDataRelease1/. We use the TEMPO2 software package (Edwards et al. 2006; Hobbs et al. 2006) to process the real and synthetic data. The software for applying the Kalman filter-based parameter estimation method discussed in this paper to real pulsar data and for carrying out simulations with this method is available at https://github.com/oneill-academic/pulsar_freq_filter.

REFERENCES

Alpar M. A., Chau H. F., Cheng K. S., Pines D., 1993, ApJ, 409, L345
Anderson P. W., Itoh N., 1975, Nature, 256, 25
Andersson N., Comer G. L., Prix R., 2003, Phys. Rev. Lett., 90, 091101
Antonelli M., Basu A., Haskell B., 2023, MNRAS, 520, 2813
Antonelli M., Montoli A., Pizzochero P. M., 2022, in Vasconcellos C. A. Z., ed., Astrophysics in the XXI Century with Compact Stars. World Scientific, Singapore, p. 219
Antonopoulou D., Haskell B., Espinoza C. M., 2022, Rep. Prog. Phys., 85, 126901
Arzoumanian Z., Nice D. J., Taylor J. H., Thorsett S. E., 1994, ApJ, 422, L671
Bailes M. et al., 2017, PASA, 34, e045
Baym G., Pethick C., Pines D., Ruderman M., 1969, Nature, 224, 872
Bellman R., Åström K., 1970, Mathematical Biosciences, 7, 329
Boynton P. E., Groth E. J., Hutchinson D. P., Nanos G. P. J., Partridge R. B., Wilkinson D. T., 1972, ApJ, 175, L217
Çerri-Serim D., Serim M. M., Şahiner Ş., İnam S. Ç., Baykal A., 2019, MNRAS, 485, 2
Chamel N., 2012, Phys. Rev. C, 85, 035801
Cheng K. S., 1987, ApJ, 321, L799
Chugunov A. I., Horowitz C. J., 2010, MNRAS, 407, 54
Coles W., Hobbs G., Champion D. J., Manchester R. N., Verbiest J. P. W., 2011, MNRAS, 418, 561
Cordes J. M., 1980, ApJ, 237, L216
Cordes J. M., Downs G. S., 1985, ApJS, 59, 343
Cordes J. M., Helfand D. J., 1980, ApJ, 239, L640
Deeter J. E., 1984, ApJ, 281, L482
Dodson R. G., McCulloch P. M., Lewis D. R., 2002, ApJ, 564, L85
Draper N. R., Smith H., 1998, Applied Regression Analysis. Third Edition. John Wiley and Sons, Inc., New York, NY, p. 115
Drummond L. V., Melatos A., 2017, MNRAS, 472, 4851
Dunn L. et al., 2022, MNRAS, 512, 1469
Edwards R. T., Hobbs G. B., Manchester R. N., 2006, MNRAS, 372, 1549
Espinoza C. M., Antonopoulou D., Dodson R., Stepanova M., Scherer A., 2021, A&A, 647, 25
Espinoza C. M., Lyne A. G., Stappers B. W., Kramer M., 2011, MNRAS, 414, 1679
Ferrari A., Ruffini R., 1969, ApJ, 158, L71
Gelb A., ed. 1974, Applied Optimal Estimation. The MIT Press, Cambridge, MA, p. 78
Giliberti E., Cambiotti G., 2022, MNRAS, 511, 3365
Goldreich P., Julian W. H., 1969, ApJ, 157, L869
Goncharov B. et al., 2021, MNRAS, 502, 478
Goncharov B., Zhu X.-J., Thrane E., 2020, MNRAS, 497, 3264
Greenstein G., 1970, Nature, 227, 791
Groth E. J., 1975, ApJS, 29, 453
Gügercinoğlu E., Alpar M. A., 2017, MNRAS, 471, 4827
Gügercinoğlu E., Ge M. Y., Yuan J. P., Zhou S. Q., 2022, MNRAS, 511, 425
Haskell B., Melatos A., 2015, Int. J. Mod. Phys. D, 24, 1530008
Helfand D. J., Taylor J. H., Backus P. R., Cordes J. M., 1980, ApJ, 237, L206
Hobbs G. B., Edwards R. T., Manchester R. N., 2006, MNRAS, 369, 655
Hobbs G., Lyne A. G., Kramer M., 2010, MNRAS, 402, 1027
Kalman R. E., 1960, Transactions of the ASME–Journal of Basic Engineering, 82, 35
Keith M. J., Niţu I. C., 2023, MNRAS, 523, 4603
Kerin A. D., Melatos A., 2022, MNRAS, 514, 1628
Lamb F. K., Pines D., Shaham J., 1978a, ApJ, 224, L969
Lamb F. K., Pines D., Shaham J., 1978b, ApJ, 225, L582
Lentati L. et al., 2016, MNRAS, 458, 2161

Lentati L., Alexander P., Hobson M. P., Feroz F., van Haasteren R., Lee K. J., Shannon R. M., 2014, MNRAS, 437, 3004
Link B., Epstein R. I., Lattimer J. M., 1999, Phys. Rev. Lett., 83, 3362
Lower M. E. et al., 2020, MNRAS, 494, 228
Lower M. E. et al., 2021, MNRAS, 508, 3251
Lyne A. G., Shemar S. L., Smith F. G., 2000, MNRAS, 315, 534
Matsakis D. N., Taylor J. H., Eubanks T. M., 1997, A&A, 326, 924
McCulloch P. M., Hamilton P. A., McConnell D., King E. A., 1990, Nature, 346, 822
Melatos A., Link B., 2014, MNRAS, 437, 21
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
Middleditch J., Marshall F. E., Wang Q. D., Gotthelf E. V., Zhang W., 2006, ApJ, 652, L1531
Ostriker J. P., Gunn J. E., 1969, ApJ, 157, L1395
Parthasarathy A. et al., 2019, MNRAS, 489, 3810
Parthasarathy A. et al., 2020, MNRAS, 494, 2012
Pines D., Alpar M. A., 1985, Nature, 316, 27
Price S., Link B., Shore S. N., Nice D. J., 2012, MNRAS, 426, 2507
Ruderman M., Zhu T., Chen K., 1998, ApJ, 492, L267
Shannon R. M., Cordes J. M., 2010, ApJ, 725, L1607
Shemar S. L., Lyne A. G., 1996, MNRAS, 282, 677
Speagle J. S., 2020, MNRAS, 493, 3132
van Eysden C. A., Melatos A., 2010, MNRAS, 409, 1253
van Haasteren R., Levin Y., 2013, MNRAS, 428, 1147
van Haasteren R., Levin Y., McDonald P., Lu T., 2009, MNRAS, 395, 1005
Vargas A. F., Melatos A., 2023, MNRAS, 522, 4880
Verbiest J. P. W., Shaifullah G. M., 2018, Class. Quantum Gravity, 35, 133001
Wang N., Manchester R. N., Pace R. T., Bailes M., Kaspi V. M., Stappers B. W., Lyne A. G., 2000, MNRAS, 317, 843
Warszawski L., Melatos A., 2011, MNRAS, 415, 1611
Wong T., Backer D. C., Lyne A. G., 2001, ApJ, 548, L447
Yu M. et al., 2013, MNRAS, 429, 688
Yuan J. P., Wang N., Manchester R. N., Liu Z. Y., 2010, MNRAS, 404, 289
Zhou S., Gügercinoğlu E., Yuan J., Ge M., Yu C., 2022, Universe, 8, 641

APPENDIX A: POWER SPECTRUM

In this appendix, we derive the PSD of the dependent variables Ω_c and Ω_s in the two-component model given by the differential equations (1) and (2). The behaviour and power spectrum of the stochastic part of the solutions are not affected by the constant torques so they are removed from the differential equations for this calculation. This is equivalent to subtracting a linear best fit and analysing the residuals. A similar calculation appears in Appendix A of Meyers et al. (2021b) and has been generalized to more than two stellar components by Antonelli et al. (2023).

Upon Fourier transforming equations (1) and (2), one obtains

$$i\omega\hat{\Omega}_c(\omega) = -\frac{1}{\tau_c}\hat{\Omega}_c(\omega) + \frac{1}{\tau_c}\hat{\Omega}_s(\omega) + \frac{\hat{\xi}_c(\omega)}{I_c}, \tag{A1}$$

$$i\omega\hat{\Omega}_s(\omega) = -\frac{1}{\tau_s}\hat{\Omega}_s(\omega) + \frac{1}{\tau_s}\hat{\Omega}_c(\omega) + \frac{\hat{\xi}_s(\omega)}{I_s}. \tag{A2}$$

Solving equations (A1) and (A2) for $\hat{\Omega}_c(\omega)$ and $\hat{\Omega}_s(\omega)$ gives

$$\hat{\Omega}_c(\omega) = \frac{\left(i\omega + \frac{1}{\tau_s}\right)\frac{\hat{\xi}_c(\omega)}{I_c} + \frac{1}{\tau_c}\frac{\hat{\xi}_s(\omega)}{I_s}}{-\omega^2 + i\omega/\tau}, \tag{A3}$$

$$\hat{\Omega}_s(\omega) = \frac{\frac{1}{\tau_s}\frac{\hat{\xi}_c(\omega)}{I_c} + \left(i\omega + \frac{1}{\tau_c}\right)\frac{\hat{\xi}_s(\omega)}{I_s}}{-\omega^2 + i\omega/\tau}. \tag{A4}$$

Given two random variables $x(t)$ and $y(t)$, their cross-PSD $P_{xy}(\omega)$ (or simply the PSD if $x = y$) can be calculated by the formula

$$\langle \hat{x}^*(\omega)\hat{y}(\omega')\rangle = 2\pi\delta(\omega - \omega')P_{xy}(\omega). \tag{A5}$$

In this paper it is assumed that ξ_c/I_c and ξ_s/I_s are uncorrelated and stationary white noise processes with flat spectra, so we have

$$P_{\xi_c\xi_c}(\omega) = \sigma_c^2, \tag{A6}$$

$$P_{\xi_s\xi_s}(\omega) = \sigma_s^2, \tag{A7}$$

$$P_{\xi_c\xi_s}(\omega) = 0. \tag{A8}$$

Combining equations (A6)–(A8) with equations (A3) and (A4) gives

$$P_{\Omega_c\Omega_c}(\omega) = \frac{\left(\omega^2 + \frac{1}{\tau_s^2}\right)\frac{\sigma_c^2}{I_c^2} + \frac{1}{\tau_c^2}\frac{\sigma_s^2}{I_s^2}}{\omega^4 + \omega^2/\tau^2}, \tag{A9}$$

$$P_{\Omega_s\Omega_s}(\omega) = \frac{\left(\omega^2 + \frac{1}{\tau_c^2}\right)\frac{\sigma_s^2}{I_s^2} + \frac{1}{\tau_s^2}\frac{\sigma_c^2}{I_c^2}}{\omega^4 + \omega^2/\tau^2}, \tag{A10}$$

$$P_{\Omega_c\Omega_s}(\omega) = \frac{i\omega\left(\frac{\sigma_s^2}{\tau_c I_s^2} - \frac{\sigma_c^2}{\tau_s I_c^2}\right) + \left(\frac{\sigma_c^2}{\tau_s^2 I_c^2} + \frac{\sigma_s^2}{\tau_c^2 I_s^2}\right)}{\omega^4 + \omega^2/\tau^2}. \tag{A11}$$

Equations (A9) and (A10) asymptote towards pure power laws as functions of ω in the limits of high and low ω. Specifically, the PSDs of Ω_c and Ω_s tend to a spectral index of -2 at high and low ω; see Meyers et al. (2021b) for more details.

In the literature, timing noise is often described by the power spectrum of timing residuals, which is often written in the form

$$P(f) = \frac{A^2}{12\pi^2 f_{yr}^3}\left(\frac{f}{f_{yr}}\right)^{\gamma} \tag{A12}$$

(Lentati et al. 2016). In the limit $\omega \gg 1/\tau$, equation (A9) implies

$$P(f) = \frac{Q_c}{(2\pi)^4\Omega_0^2 f^4}, \tag{A13}$$

noting that $P(f) = P_{\Omega_c\Omega_c}(\omega)/(\Omega_0^2\omega^2)$, where $\Omega_0 = 49.28$ rad s^{-1} is the pulsar's angular frequency. If we set $\gamma = -4$ in equation (A12) we can convert Q_c to A^2 by the formula

$$A^2 = \frac{12\pi^2}{(2\pi)^4\Omega_0^2 f_{yr}} Q_c. \tag{A14}$$

Numerically we have

$$\log_{10} A = 0.5(\log_{10} Q_c + 2.99). \tag{A15}$$

In equation (A15) and elsewhere, A is unitless and Q_c has units of rad^{2}s^{-3}.

APPENDIX B: STATE SPACE REPRESENTATION

In this appendix, we give the full forms of the Kalman filter update matrices $\boldsymbol{F}_i$, $\boldsymbol{T}_i$, and $\boldsymbol{Q}_i$ defined in equations (6) and (8) to assist the interested reader in reproducing the numerical results in Section 5. Specifically, we have

$$\boldsymbol{F}_i = \frac{1}{\tau_s+\tau_c}\begin{pmatrix} \tau_c + \tau_s e^{-\Delta t_i/\tau} & \tau_s - \tau_s e^{-\Delta t_i/\tau} \\ \tau_c - \tau_c e^{-\Delta t_i/\tau} & \tau_s + \tau_c e^{-\Delta t_i/\tau} \end{pmatrix}, \tag{B1}$$

$$\boldsymbol{N}_i = \langle\dot{\Omega}_c\rangle\Delta t_i\begin{pmatrix}1\\1\end{pmatrix} + \langle\Omega_c - \Omega_s\rangle\left(1 - e^{-\Delta t_i/\tau}\right)\begin{pmatrix}\tau/\tau_c\\ -\tau/\tau_s\end{pmatrix}, \tag{B2}$$

$$\boldsymbol{Q}_i = \left(\frac{1}{\tau_c+\tau_s}\right)^2\begin{pmatrix} a & b \\ b & c\end{pmatrix}, \tag{B3}$$

with

$$a = \left(\frac{\sigma_c^2\tau_c^2}{I_c^2} + \frac{\sigma_s^2\tau_s^2}{I_s^2}\right)\Delta t_i + 2\tau\tau_s\left(\frac{\sigma_c^2\tau_c}{I_c^2} - \frac{\sigma_s^2\tau_s}{I_s^2}\right)\left(1 - e^{-\Delta t_i/\tau}\right) + \frac{\tau\tau_s^2}{2}\left(\frac{\sigma_c^2}{I_c^2} + \frac{\sigma_s^2}{I_s^2}\right)\left(1 - e^{-2\Delta t_i/\tau}\right), \tag{B4}$$

$$b = \left(\frac{\sigma_c^2\tau_c^2}{I_c^2} + \frac{\sigma_s^2\tau_s^2}{I_s^2}\right)\Delta t_i + \tau(\tau_s - \tau_c)\left(\frac{\sigma_c^2\tau_c}{I_c^2} - \frac{\sigma_s^2\tau_s}{I_s^2}\right)\left(1 - e^{-\Delta t_i/\tau}\right) - \frac{\tau\tau_c\tau_s}{2}\left(\frac{\sigma_c^2}{I_c^2} + \frac{\sigma_s^2}{I_s^2}\right)\left(1 - e^{-2\Delta t_i/\tau}\right), \tag{B5}$$

$$c = \left(\frac{\sigma_c^2\tau_c^2}{I_c^2} + \frac{\sigma_s^2\tau_s^2}{I_s^2}\right)\Delta t_i + 2\tau\tau_c\left(\frac{\sigma_s^2\tau_s}{I_s^2} - \frac{\sigma_c^2\tau_c}{I_c^2}\right)\left(1 - e^{-\Delta t_i/\tau}\right) + \frac{\tau\tau_c^2}{2}\left(\frac{\sigma_c^2}{I_c^2} + \frac{\sigma_s^2}{I_s^2}\right)\left(1 - e^{-2\Delta t_i/\tau}\right). \tag{B6}$$

APPENDIX C: KALMAN FILTER

In this appendix, for the sake of completeness and reproducibility, we summarize the structure and operation of the Kalman filter used to generate the results in Sections 5 and 6. The discussion includes a short justification of the form of the Kalman likelihood in equation (18).

The Kalman filter is a predictor–corrector algorithm. Given a sequence of $i-1$ measurements, $\boldsymbol{Y}_1, \boldsymbol{Y}_2, ..., \boldsymbol{Y}_{i-1}$, the Kalman filter predicts the value of the state $\boldsymbol{X}_i$ at the next time-step t_i. The estimate is denoted by $\boldsymbol{X}_i^{i-1}$, and its error is given by the covariance matrix $\boldsymbol{P}_i^{i-1}$. When the measurement $\boldsymbol{Y}_i$ is made, the estimate $\boldsymbol{X}_i^{i-1}$ is updated to give the new estimate of the current state, $\boldsymbol{X}_i^{i}$, and its error, $\boldsymbol{P}_i^{i}$. This procedure is carried out from t_1 to t_N.

The following update formulas, equations (C1)–(C9), implement the prediction and correction steps described earlier. The symbols are explained in the text immediately following.

Initialization:

$$\boldsymbol{X}_0^0 = \boldsymbol{\mu}_0, \tag{C1}$$

$$\boldsymbol{P}_0^0 = \boldsymbol{S}_0. \tag{C2}$$

State prediction:

$$\boldsymbol{X}_i^{i-1} = \boldsymbol{F}_i\boldsymbol{X}_{i-1}^{i-1} + \boldsymbol{T}_i, \tag{C3}$$

$$\boldsymbol{P}_i^{i-1} = \boldsymbol{F}_i\boldsymbol{P}_{i-1}^{i-1}\boldsymbol{F}_i^T + \boldsymbol{Q}_i. \tag{C4}$$

State correction:

$$\boldsymbol{\epsilon}_i = \boldsymbol{Y}_i - E(\boldsymbol{Y}_i|\boldsymbol{Y}_{1:i-1}) = \boldsymbol{Y}_i - \boldsymbol{C}\boldsymbol{X}_i^{i-1}, \tag{C5}$$

$$\boldsymbol{S}_i = \boldsymbol{C}\boldsymbol{P}_i^{i-1}\boldsymbol{C}^T + \boldsymbol{R}_i, \tag{C6}$$

$$\boldsymbol{K}_i = \boldsymbol{P}_i^{i-1}\boldsymbol{C}^T\boldsymbol{S}_i^{-1}, \tag{C7}$$

$$\boldsymbol{X}_i^{i} = \boldsymbol{X}_i^{i-1} + \boldsymbol{K}_i\boldsymbol{\epsilon}_i, \tag{C8}$$

$$\boldsymbol{P}_i^{i} = (1 - \boldsymbol{K}_i\boldsymbol{C})\boldsymbol{P}_i^{i-1}. \tag{C9}$$

Equations (C1) and (C2) estimate the initial state and its error. Equations (C3) and (C4) predict the next state from the previous data. Equations (C5)–(C9) combine the new measurement with the prediction to get a new estimate of the current state. In equation (C5), we define $\boldsymbol{\epsilon}_i = \boldsymbol{Y}_i - E(\boldsymbol{Y}_i|\boldsymbol{Y}_{1:i-1})$. In equation (C6), we define $\boldsymbol{S}_i = \mathrm{var}(\boldsymbol{\epsilon}_i)$. The matrix $\boldsymbol{K}_i$ in equations (C7)–(C9) is usually called the Kalman gain.

To get the likelihood, $p(\boldsymbol{Y} = \boldsymbol{Y}_{1:N}|\boldsymbol{\theta})$, from the Kalman filter estimates, we apply standard rules for conditional probability to get

$$p(\boldsymbol{Y}_{1:N}|\boldsymbol{\theta}) = p(\boldsymbol{Y}_N|\boldsymbol{Y}_{1:N-1}, \boldsymbol{\theta})p(\boldsymbol{Y}_{1:N-1}|\boldsymbol{\theta}) \tag{C10}$$

$$= p(\boldsymbol{Y}_N|\boldsymbol{Y}_{1:N-1}, \boldsymbol{\theta})p(\boldsymbol{Y}_{N-1}|\boldsymbol{Y}_{1:N-2}, \boldsymbol{\theta})p(\boldsymbol{Y}_{1:N-2}|\boldsymbol{\theta}) \tag{C11}$$

and hence recursively

$$p(\boldsymbol{Y}_{1:N}|\boldsymbol{\theta}) = \prod_{i=1}^{N} p(\boldsymbol{Y}_i|\boldsymbol{Y}_{1:i-1}, \boldsymbol{\theta}). \tag{C12}$$

$p(\boldsymbol{Y}_i|\boldsymbol{Y}_{1:i-1}, \boldsymbol{\theta})$ is Gaussian if all the errors are assumed to be Gaussian. The mean and covariance matrix of the $p(\boldsymbol{Y}_{1:N}|\boldsymbol{\theta})$ distribution are

$$E(\boldsymbol{Y}_i|\boldsymbol{Y}_{1:i-1}) = \boldsymbol{C}\boldsymbol{X}_i^{i-1}, \tag{C13}$$

$$\mathrm{var}(\boldsymbol{Y}_i|\boldsymbol{Y}_{1:i-1}) = \boldsymbol{S}_i, \tag{C14}$$

which imply

$$p(\boldsymbol{Y}_i|\boldsymbol{Y}_{1:i-1}, \boldsymbol{\theta}) = \mathcal{N}(\boldsymbol{Y}_i; \boldsymbol{C}\boldsymbol{X}_i^{i-1}, \boldsymbol{S}_i), \tag{C15}$$

where $\mathcal{N}(\boldsymbol{X}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes a Gaussian with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Equivalently, writing $\boldsymbol{\epsilon}_i = \boldsymbol{Y}_i - \boldsymbol{C}\boldsymbol{X}_i^{i-1}$, we obtain

$$p(\boldsymbol{Y}_i|\boldsymbol{Y}_{1:i-1}, \boldsymbol{\theta}) = p(\boldsymbol{\epsilon}_i|\boldsymbol{Y}_{1:i-1}, \boldsymbol{\theta}) \tag{C16}$$

$$= \mathcal{N}(\boldsymbol{\epsilon}_i; 0, \boldsymbol{S}_i). \tag{C17}$$

The full formula for the log-likelihood then becomes

$$\log p(\boldsymbol{Y}_{1:N}|\boldsymbol{\theta}) = \sum_{i=1}^{N} \log p(\boldsymbol{Y}_i|\boldsymbol{Y}_{1:i-1}, \boldsymbol{\theta}) \tag{C18}$$

$$= -\frac{1}{2}\sum_{i=1}^{N}\left[N_{\boldsymbol{Y}}\log(2\pi) + \log\det(\boldsymbol{S}_i) + \boldsymbol{\epsilon}_i^T\boldsymbol{S}_i^{-1}\boldsymbol{\epsilon}_i\right], \tag{C19}$$

where $N_{\boldsymbol{Y}}$ is the dimension of $\boldsymbol{Y}$. Equation (C18) follows from equation (C12) and equation (C19) follows from equation (C17). Equation (C19) is the same as equation (18) in subsection 3.3.

Once $p(\boldsymbol{Y}_{1:N}|\boldsymbol{\theta})$ is known, Bayes's theorem can be used to get the probability of the parameters in terms of the data

$$p(\boldsymbol{\theta}|\boldsymbol{Y}_{1:N}) = \frac{p(\boldsymbol{Y}_{1:N}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\boldsymbol{Y}_{1:N})}. \tag{C20}$$

APPENDIX D: PARAMETER ESTIMATION WITH SIMULATED FREQUENCIES

D1 Simulation under ideal conditions

In this appendix, our parameter estimation method is tested on simulated frequency data. To simulate the data, we must first select values for the model parameters τ_c, τ_s, σ_c/I_c, σ_s/I_s, N_c/I_c, N_s/I_s. We then chose an initial value $\Omega_c(0)$ and set $\Omega_s(0)$ to be

$$\Omega_s(0) = \Omega_c(0) - \langle\Omega_c - \Omega_s\rangle \tag{D1}$$

$$= \Omega_c(0) - \tau\left(\frac{N_c}{I_c} - \frac{N_s}{I_s}\right) \tag{D2}$$

so the pulsar is initially in equilibrium. We choose $N_{\rm obs}$ random times from $t = 0$ to $t = T_{\rm obs}$, where simulated measurements occur. The differential equations (1) and (2) are then integrated from the initial time to each measurement time, giving $\Omega_{\rm c}(t_i)$ and $\Omega_{\rm s}(t_i)$ values at each $1 \le i \le N_{\rm obs}$. Realistic timing experiments only yield $\Omega_{\rm c}$ observations, so the $\Omega_{\rm s}$ values are discarded. To simulate measurement errors we add a number drawn from a Gaussian with variance R to each $\Omega_{\rm c}$ value. For simplicity in these simulations we assume every data point has the same measurement uncertainty, unlike in real data. Once the data are simulated, the parameter estimation algorithm is executed, and the posterior distribution is compared to the true parameters.

Fig. D1 shows the results for 200 simulations with randomly chosen values for r, τ, $Q_{\rm c}$, $Q_{\rm s}$, $\langle \Omega_{\rm c} - \Omega_{\rm s} \rangle$, $\langle \dot{\Omega}_{\rm c} \rangle$ and with $\Omega_{\rm c}(0) = 10$ rad s^{-1}. The simulations are under idealized conditions (i.e. low noise, many samples) with $T_{\rm obs} = 1000$ d, $N_{\rm obs} = 1000$, and $R = 10^{-30}$ rad^{2}s^{-2}. The six panels plot the recovered values for each of the six parameters against the injected parameters. The diagonal red lines mark where the recovered and injected parameters are equal, indicating a successful simulation. The colours indicate the width of the marginalized posterior for that parameter (width here meaning the interquartile range). The closeness of the points to the red line in the τ, $Q_{\rm c}$, and $\langle \dot{\Omega}_{\rm c} \rangle$ panels means those parameters are generally well recovered, while the large vertical spread of the points in the r, $Q_{\rm s}$, and $\langle \Omega_{\rm c} - \Omega_{\rm s} \rangle$ panels means that they are usually harder to recover. This agrees with the identifiability analyses in subsection 3.2 and Appendix E. Interestingly, small τ values tend to be better recovered, which is unsurprising since these correspond to strong damping.

As a further verification of the method, in Fig. D2 we show a standard PP-plot (probability–probability plot) for the same 200 simulations as in Fig. D1. The curves for each parameter remain close to the diagonal line, indicating that the posteriors give unbiased estimates. For more details on interpretation of PP-plots see Meyers et al. (2021b).

D2 Measurement noise

In this appendix, we test the Kalman filter's ability to recover parameters with different levels of measurement error. Large measurement errors make it difficult for the Kalman filter to isolate the true random process, making the parameters harder to recover. To show the effect of large measurement noise we run three sets of 200 simulations, with $R = 10^{-24}$ rad^{2}s^{-2}, 10^{-20} rad^{2}s^{-2} and 10^{-16} rad^{2}s^{-2}. The results are shown in the three columns of Fig. D3. It is easiest to see the effect of varying R by looking at τ and $Q_{\rm c}$ because they are the best recovered (except $\langle \dot{\Omega}_{\rm c} \rangle$ which is trivial). We can see in each test that $Q_{\rm c}$ is poorly recovered below some critical value, when the measurement noise makes up most of the total noise. If $Q_{\rm c}$ is above the critical value, parameter estimation is usually successful.

D3 Impact of incorrect measurement uncertainties

In this section, we investigate the effect of feeding incorrect measurement uncertainties into the Kalman filter and introduce a modified parameter estimation algorithm to correct for the effect.

The frequency uncertainties for the real data in this paper are calculated from the TOA uncertainties provided in the UTMOST data release (Lower et al. 2020). Frequency uncertainties may be wrong, if the raw TOA errors or their conversion to frequency errors are incorrect. The quoted TOA uncertainties depend on many factors such as dispersion measure (Verbiest & Shaifullah 2018).

To simulate the problem, we generate data with a measurement error variance $R_{\rm true}$ but use a different measurement error variance, $R_{\rm KF}$, in the Kalman filter. Results with $R_{\rm KF} = 10^{-23}$ rad^{2}s^{-2}, $R_{\rm true} = 10^{-27}$ rad^{2}s^{-2} and $R_{\rm KF} = 10^{-27}$ rad^{2}s^{-2}, $R_{\rm true} = 10^{-23}$ rad^{2}s^{-2} are shown in the left hand columns of Fig. D4 and Fig. D5, respectively. The results suggest that when $R_{\rm KF}$ is too large (column 1 of Fig. D4) the $Q_{\rm c}$ values are underestimated and when $R_{\rm KF}$ is too small (column 1 of Fig. D5) the $Q_{\rm c}$ values are overestimated. The effect is most pronounced for $Q_{\rm c} \lesssim 10^{-26}$ rad^{2}s^{-3}.

The Kalman filter separates measurement noise from the underlying random process guided by $R_{\rm KF}$. If it is guided to remove too much noise ($R_{\rm KF} \ge R_{\rm true}$), then some process noise is removed too, and the inferred strength of the process noise is weaker than it should be. The reverse is true for $R_{\rm KF} \le R_{\rm true}$. If the measurement noise is not properly separated then τ is also difficult to recover.

We adjust for $R_{\rm KF}$ being input incorrectly by sampling R as well as the six dynamical model parameters r, τ, $Q_{\rm c}$, $Q_{\rm s}$, $\langle \Omega_{\rm c} - \Omega_{\rm s} \rangle$, $\langle \dot{\Omega}_{\rm c} \rangle$. The results are shown in the right columns of Fig. D4 and Fig. D5. The recovered $Q_{\rm c}$ values are now close to their correct values and the spread on the recovered τ values is lower, which is an encouraging outcome. When analysing the real data in Section 5, sampling R makes little difference to the posterior, so the result is not displayed for brevity. We present this modified estimation method in this appendix in preparation for analysing more objects in the future.

APPENDIX E: IDENTIFIABILITY OF NOISE PARAMETERS

In this appendix, we consider a simplified version of the parameter estimation problem to get a heuristic for the identifiability of the noise parameters $Q_{\rm c}$ and $Q_{\rm s}$, as we did in subsection 3.2 for the other parameters. We aim to separate the measurement noise from $\xi_{\rm c}$ and $\xi_{\rm s}$ by exploiting their different behaviours over time. Specifically, $\xi_{\rm c}$ acts directly on the crust so it affects the frequency faster than $\xi_{\rm s}$, whose effect is delayed by the coupling time-scale. The measurement noise does not grow with time, nor does it evolve according to the equations of motion. The relative strengths of ξ_c, $\xi_{\rm s}$, and the measurement noise affect the size and shape of the observed random walk in $\Omega_{\rm c}$. By looking at the rate that the random walk in $\Omega_{\rm c}$ grows over different time-scales, the strengths of the three noise types can be estimated. In this section we set $N_{\rm c} = N_{\rm s} = 0$ to analyse just the random behaviour.

Let $\Omega_{\rm c}$ be observed at some instant and then again a time Δt later, such that $\Omega_{\rm c}$ changes by an amount $\Delta\Omega_{\rm c}$. Because it is a random process, $\Delta\Omega_{\rm c}$ is drawn from a random distribution with a variance we can calculate as a function of Δt. By comparing the observed distribution of $\Delta\Omega_{\rm c}$ to the calculated distribution we can fit for the unknown parameters to infer $Q_{\rm c}$, $Q_{\rm s}$, and the size of the measurement noise.

The variance of a jump in $\Omega_{\rm c}$ over a time Δt is given by equation (B3), viz.

$$Q_{0,0}(\Delta t) = \langle (\Omega_{\rm c}(t+\Delta t) - \Omega_{\rm c}(t))^2 \rangle \tag{E1}$$

$$= \frac{Q_{\rm c}\tau_{\rm c}^2 + Q_{\rm s}\tau_{\rm s}^2}{(\tau_{\rm c}+\tau_{\rm s})^2}\Delta t \tag{E2}$$

$$+ 2\tau\tau_{\rm s}\frac{Q_{\rm c}\tau_{\rm c} - Q_{\rm s}\tau_{\rm s}}{(\tau_{\rm c}+\tau_{\rm s})^2}\left(1 - e^{-\Delta t/\tau}\right)$$

$$+ \frac{\tau\tau_{\rm s}^2}{2}\frac{Q_{\rm c}+Q_{\rm s}}{(\tau_{\rm c}+\tau_{\rm s})^2}\left(1 - e^{-2\Delta t/\tau}\right)$$

$$= Q_{\rm c}T_{\rm c}(\Delta t) + Q_{\rm s}T_{\rm s}(\Delta t), \tag{E3}$$

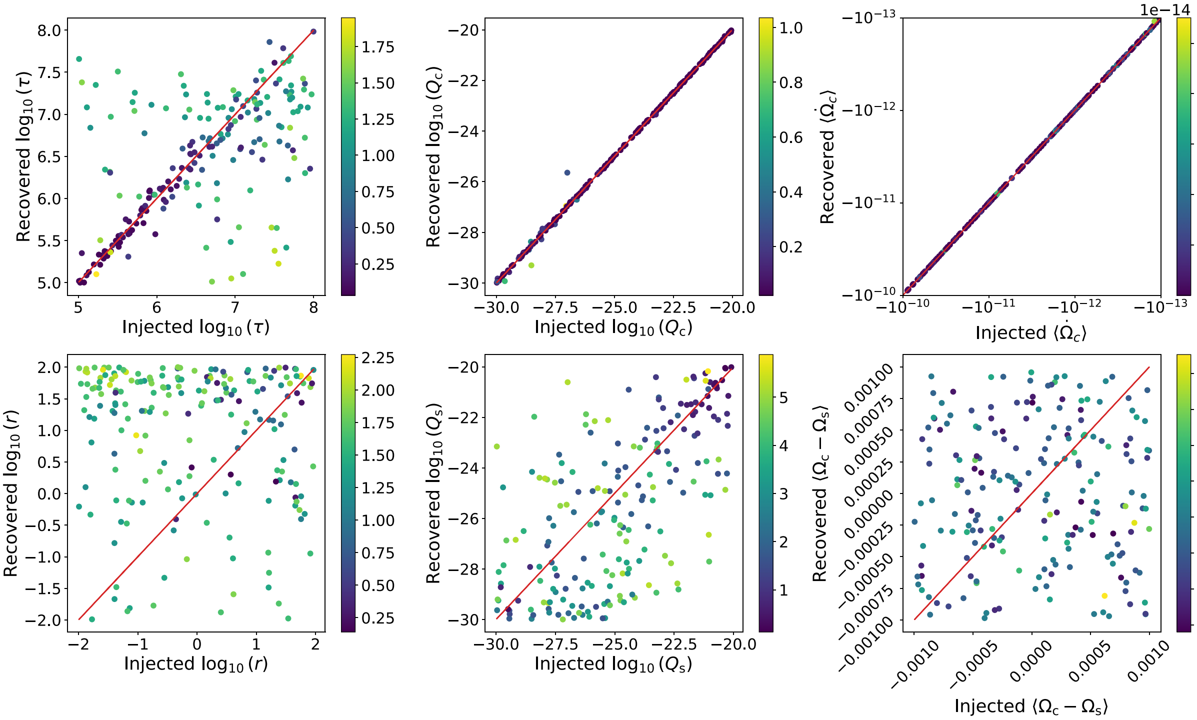

Figure D1. Test of the parameter estimation scheme on simulated frequency data showing the injected (horizontal axis) and recovered (vertical axis) r, τ, Q_c, Q_s, $\langle\Omega_c - \Omega_s\rangle$, and $\langle\dot{\Omega}_c\rangle$ values for 200 simulations. The injected parameters are those used to simulate the data and the recovered parameters are the peak of the posterior distribution obtained using the Kalman filter and nested sampler. The colour of each point can be compared to the colour bar next to its panel to get the width (interquartile range) of the marginalized posterior for that parameter.

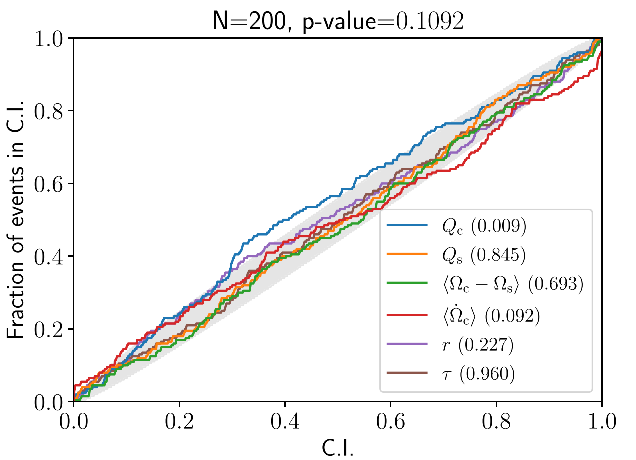

Figure D2. PP-plot for the 200 simulations in Fig. D1, with the confidence interval on the horizontal axis and the proportion of simulations where the true values lie within that confidence interval on the vertical axis. The posterior distributions produced by the procedure in this paper are reliable if the curves stay close to the diagonal line. Specifically, if they stay within the shaded region then one can be 90 per cent confident that the posteriors are unbiased. For each parameter the legend lists the colour of the curve corresponding to it and p-values for the distribution of that parameter derived from a Kolmogorov–Smirnov test.

with

$$T_c(\Delta t) = \frac{\tau_c^2 \Delta t + 2\tau\tau_c\tau_s\left(1 - e^{-\Delta t/\tau}\right) + \frac{\tau\tau_s^2}{2}\left(1 - e^{-2\Delta t/\tau}\right)}{(\tau_c + \tau_s)^2} \tag{E4}$$

and

$$T_s(\Delta t) = \frac{\tau_s^2 \Delta t - 2\tau\tau_s^2\left(1 - e^{-\Delta t/\tau}\right) + \frac{\tau\tau_s^2}{2}\left(1 - e^{-2\Delta t/\tau}\right)}{(\tau_c + \tau_s)^2}. \tag{E5}$$

For $\Delta t \gg \tau$ one has

$$Q_{0,0} \approx \frac{Q_c\tau_c^2 + Q_s\tau_s^2}{(\tau_c + \tau_s)^2}\Delta t \tag{E6}$$

and for $\Delta t \ll \tau$ one has

$$Q_{0,0} \approx Q_c \Delta t. \tag{E7}$$

Hence the influence of ξ_s through Q_s on short-time-scales is negligible.

The variance for a step in the measurement can be calculated from equation (E3). Denote the measurement at time t_i by Y_i and the measurement noise by v_i. Then one has

$$Y_i = \Omega_{c,i} + v_i. \tag{E8}$$

In this appendix we assume for simplicity that all the measurement errors have variance R even though in reality different data points have different measurement uncertainties. Then the change in Y between t_i and t_j has variance

$$\mathrm{var}(\Delta Y) = \mathrm{var}(Y_j - Y_i) \tag{E9}$$

$$= \mathrm{var}[(\Omega_{c,j} + v_j) - (\Omega_{c,i} + v_i)] \tag{E10}$$

$$= Q_c T_c(\Delta t) + Q_s T_s(\Delta t) + 2R. \tag{E11}$$

Let us define the auxiliary function

$$Q(\Delta t) = Q_c T_c(\Delta t) + Q_s T_s(\Delta t) + 2R. \tag{E12}$$

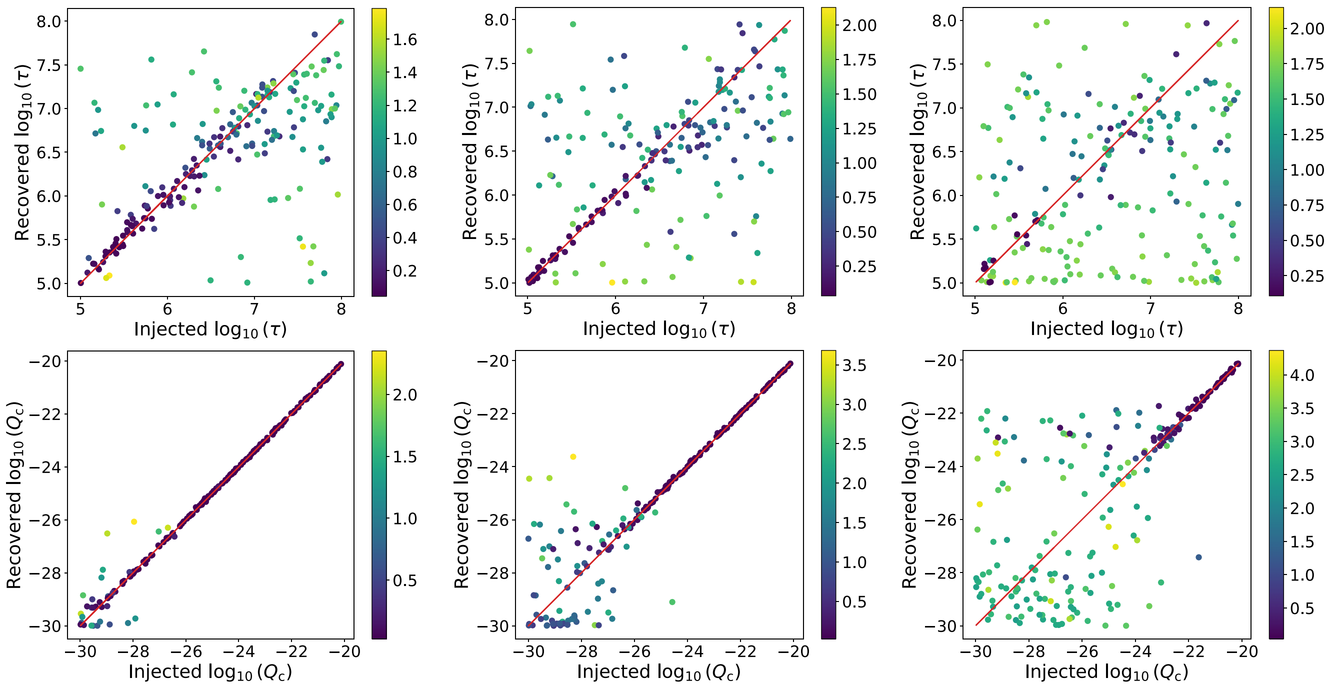

Figure D3. Test of the parameter estimation scheme on simulated frequency data showing the injected (horizontal axis) and recovered (vertical axis) τ and Q_c values (top and bottom rows, respectively). Each column shows the results of 200 simulations with a different level of measurement error. Column 1 has $R = 10^{-24}$ rad^{2}s^{-2}, column 2 has $R = 10^{-20}$ rad^{2}s^{-2}, and column 3 has $R = 10^{-16}$ rad^{2}s^{-2}. The colour of each point can be compared to the colour bar next to its panel to get the width (interquartile range) of the marginalized posterior for that parameter.

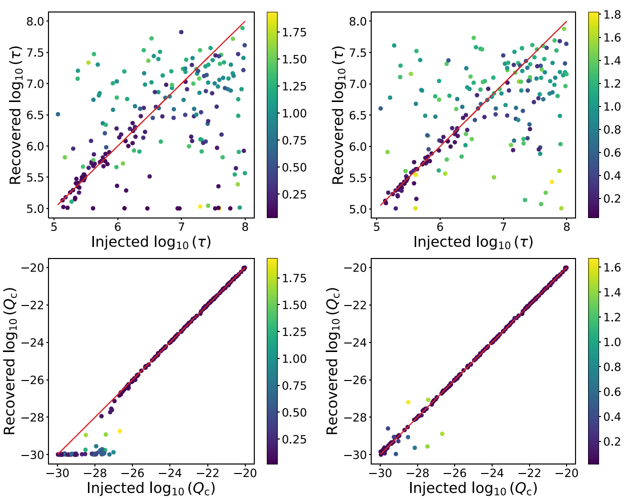

Figure D4. Test of the parameter estimation scheme on simulated frequency data showing the injected (horizontal axis) and recovered (vertical axis) τ and Q_c values (top and bottom rows, respectively) for 200 simulations with $R_{\rm true} = 10^{-27}$ rad^{2}s^{-2}, $R_{\rm KF} = 10^{-23}$ rad^{2}s^{-2}. The colour of each point can be compared to the colour bar next to its panel to get the width (interquartile range) of the marginalized posterior for that parameter. The two plots in the left-hand column show the results with the normal parameter estimation algorithm. The plots in the right-hand column show the results on the same data supplemented by sampling of $R_{\rm KF}$.

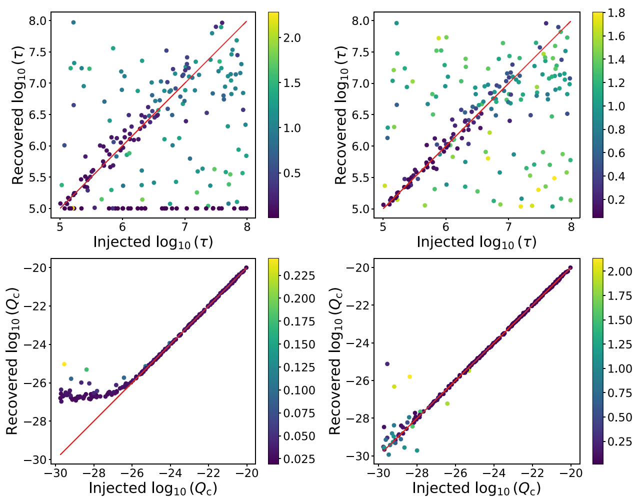

Figure D5. As for figure D4, but with $R_{\rm true} = 10^{-23}$ rad^{2}s^{-2} and $R_{\rm KF} = 10^{-27}$ rad^{2}s^{-2}.

Then $Y_j - Y_i$ is drawn from a normal distribution with variance $Q_{i,j} = Q(\Delta t_{i,j})$, with $\Delta t_{i,j} = t_j - t_i$. Calculating $\tilde{Q}_{i,j} = (Y_j - Y_i)^2$ gives a single-sample estimate of the variance function $Q_{i,j}$. The average of many $\tilde{Q}_{i,j}$'s with the same Δt converges to $Q(\Delta t)$. Let us calculate $\tilde{Q}_{i,j} = (Y_j - Y_i)^2$ for the one-step transitions, $t_1 \to t_2$, $t_2 \to t_3$,..., $t_{n-1} \to t_n$, two-step transitions, $t_1 \to t_3$,..., $t_{n-2} \to t_n$, all the way to the last and longest transition $t_1 \to t_n$. This makes $n(n-1)/2$ transitions. Let the estimate $\tilde{Q}_{i,j}$ differ from the true variance of the jump distribution, $Q_{i,j}$, by $\epsilon_{i,j}$, viz.

$$\tilde{Q}_{i,j} = Q_{i,j} + \epsilon_{i,j}. \tag{E13}$$

Then the set of equations for all the transitions is

$$\tilde{Q}_{1,2} = Q_c T_c(\Delta t_{1,2}) + Q_s T_s(\Delta t_{1,2}) + 2R + \epsilon_{1,2} \tag{E14}$$

$$\tilde{Q}_{2,3} = Q_{\rm c}T_{\rm c}(\Delta t_{2,3}) + Q_{\rm s}T_{\rm s}(\Delta t_{2,3}) + 2R + \epsilon_{2,3} \quad \text{(E15)}$$

$$\vdots$$

$$\tilde{Q}_{1,n} = Q_{\rm c}T_{\rm c}(\Delta t_{1,n}) + Q_{\rm s}T_{\rm s}(\Delta t_{1,n}) + 2R + \epsilon_{1,n}. \quad \text{(E16)}$$

If we assume we know $\tau_{\rm c}$ and $\tau_{\rm s}$ then we can calculate $T_{\rm c}$ and $T_{\rm s}$ using equations (E4) and (E5). Equations (E14)–(E16) are a regression problem for the dependent variable $\tilde{Q}$ in terms of independent variables $T_{\rm c}$ and $T_{\rm s}$ where we fit for $Q_{\rm c}$, $Q_{\rm s}$, and R. We make the simplification that the ϵ's are all drawn from the same distribution (this is not related to the R's all being the same). Standard methods for least-squares problems (Draper & Smith 1998) yield

$$\begin{pmatrix} Q_{\rm c} \\ Q_{\rm s} \\ 2R \end{pmatrix} = \begin{pmatrix} \langle T_{\rm c}^2\rangle & \langle T_{\rm c}T_{\rm s}\rangle & \langle T_{\rm c}\rangle \\ \langle T_{\rm c}T_{\rm s}\rangle & \langle T_{\rm s}^2\rangle & \langle T_{\rm s}\rangle \\ \langle T_{\rm c}\rangle & \langle T_{\rm s}\rangle & 1 \end{pmatrix}^{-1} \begin{pmatrix} \langle T_{\rm c}\tilde{Q}\rangle \\ \langle T_{\rm s}\tilde{Q}\rangle \\ \langle \tilde{Q}\rangle \end{pmatrix}, \quad \text{(E17)}$$

with

$$\mathrm{var}(Q_{\rm c}) = \frac{\langle T_{\rm s}^2\rangle - \langle T_{\rm s}\rangle^2}{D}\mathrm{var}(\epsilon), \quad \text{(E18)}$$

$$\mathrm{var}(Q_{\rm s}) = \frac{\langle T_{\rm c}^2\rangle - \langle T_{\rm c}\rangle^2}{D}\mathrm{var}(\epsilon), \quad \text{(E19)}$$

$$\mathrm{var}(R) = \frac{1}{4}\frac{\langle T_{\rm c}^2\rangle\langle T_{\rm s}^2\rangle - \langle T_{\rm c}T_{\rm s}\rangle^2}{D}\mathrm{var}(\epsilon), \quad \text{(E20)}$$

where D is the determinant of the matrix being inverted,

$$D = (\langle T_{\rm c}^2\rangle - \langle T_{\rm c}\rangle^2)(\langle T_{\rm s}^2\rangle - \langle T_{\rm s}\rangle^2) - (\langle T_{\rm c}T_{\rm s}\rangle - \langle T_{\rm c}\rangle\langle T_{\rm s}\rangle)^2. \quad \text{(E21)}$$

In equations (E18) and (E19), $\mathrm{var}(Q_{\rm c})$ has $T_{\rm s}^2$ in the numerator, whereas $Q_{\rm s}$ has $T_{\rm c}^2$. We note that $T_{\rm c}$ is larger than $T_{\rm s}$ for small time-steps. There are more one-step transitions than two-step transitions, more two-step than three-step transitions and so on. Hence we expect $\mathrm{var}(Q_{\rm c})$ to be smaller than $\mathrm{var}(Q_{\rm s})$ and $Q_{\rm c}$ should be easier to estimate from the data. Indeed, we do find that $Q_{\rm s}$ is difficult to identify in simulations, as discussed in Section D1. The above calculation justifies the empirical results logically.

The results of the simulations shown in Fig. D1 show the relative identifiabilities of the six model parameters including the two process noise parameters $Q_{\rm c}$ and $Q_{\rm s}$. The recovered $Q_{\rm c}$ points lie close to the red line, indicating that they are usually recovered accurately. The recovered $Q_{\rm s}$ points are scattered over a wider range and are not recovered as well as $Q_{\rm c}$.

APPENDIX F: TEMPO2 FITS TO LOCAL FREQUENCIES

In this appendix, we apply the parameter estimation scheme to frequencies fitted by TEMPO2 to simulated TOAs. We confirm that the low-pass filtering action of local frequency fitting introduces modest biases into the estimated parameters, chiefly $Q_{\rm c}$.

To simulate pulsar TOAs, we generate $\Omega_{\rm c}$ and $\Omega_{\rm s}$ data as in Appendix D but append a differential equation for the crust phase $\phi_{\rm c}$ to get

$$\frac{d\Omega_{\rm c}}{dt} = -\frac{1}{\tau_{\rm c}}(\Omega_{\rm c} - \Omega_{\rm s}) + \frac{N_{\rm c}}{I_{\rm c}} + \frac{\xi_{\rm c}}{I_{\rm c}} \quad \text{(F1)}$$

$$\frac{d\Omega_{\rm s}}{dt} = -\frac{1}{\tau_{\rm s}}(\Omega_{\rm s} - \Omega_{\rm c}) + \frac{N_{\rm s}}{I_{\rm s}} + \frac{\xi_{\rm s}}{I_{\rm s}} \quad \text{(F2)}$$

$$\frac{d\phi_{\rm c}}{dt} = \Omega_{\rm c}. \quad \text{(F3)}$$

A TOA occurs when the pulsar beam points towards the Earth. We assume the beam is attached rigidly to the crust, so we define $\phi_{\rm c} =$

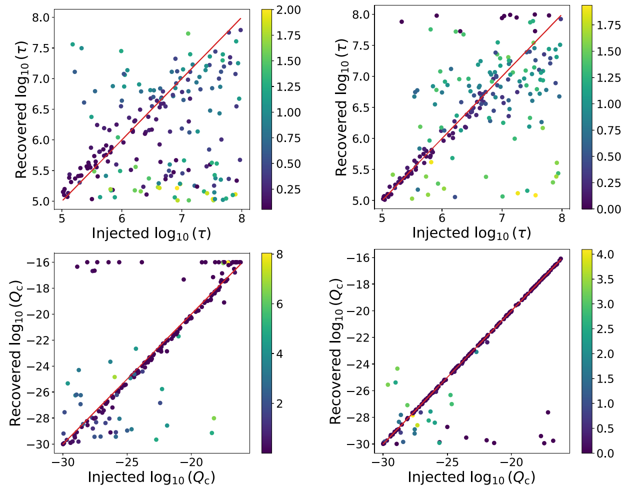

Figure F1. Accuracy of TEMPO2 fits to local frequencies. Plots of the injected (horizontal axis) and recovered (vertical axis) τ and $Q_{\rm c}$ values (top and bottom rows, respectively) for 200 simulations. The two plots in the left-hand column show the results when the method is applied to frequencies fitted to simulated TOAs with random spacings. The two plots in the right-hand column show the results when the method is applied to simulated frequencies without any TEMPO2 fitting. The colour of each point can be compared to the colour bar next to its panel to get the width (interquartile range) of the marginalized posterior for that parameter.

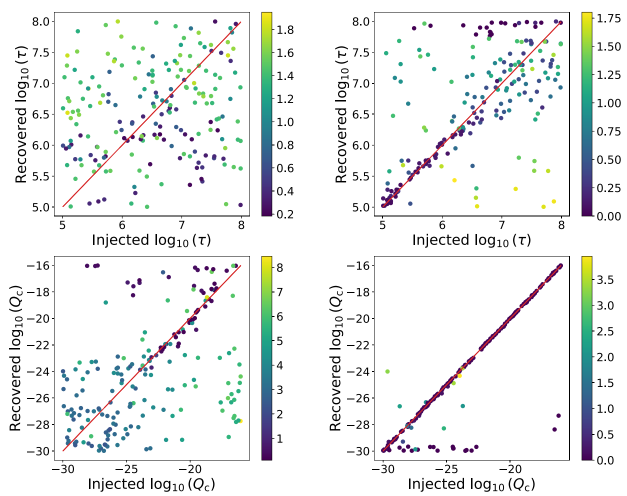

Figure F2. Accuracy of TEMPO2 fits to local frequencies for synthetic data mimicking UTMOST observations of PSR J1359−6038. Plots of the injected (horizontal axis) and recovered (vertical axis) τ and $Q_{\rm c}$ values (top and bottom rows, respectively) for 200 simulations. The two plots in the left-hand column show the results when the method is applied to frequencies fitted to simulated TOAs with the same spacing as the real data from PSR J1359−6038 used in Section 5. The two plots in the right-hand column show the results when the method is applied to simulated frequencies with the same spacing and uncertainties as on the left, but without any TEMPO2 fitting. The colour of each point can be compared to the colour bar next to its panel to get the width (interquartile range) of the marginalized posterior for that parameter.

0 (mod 2π) at a TOA. We then choose a set of observation epochs t_1, t_2,..., t_n, which we anticipate adjusting slightly to coincide with the nearest TOAs. We integrate equations (F1)–(F3) from t_1 [with $\phi_{\rm c}(t_1) = 0$] to t_2. We extrapolate linearly to find the next nearest instant $t_2' > t_2$, that gives $\phi_{\rm c}(t_2') = 0$, noting that linear extrapolation

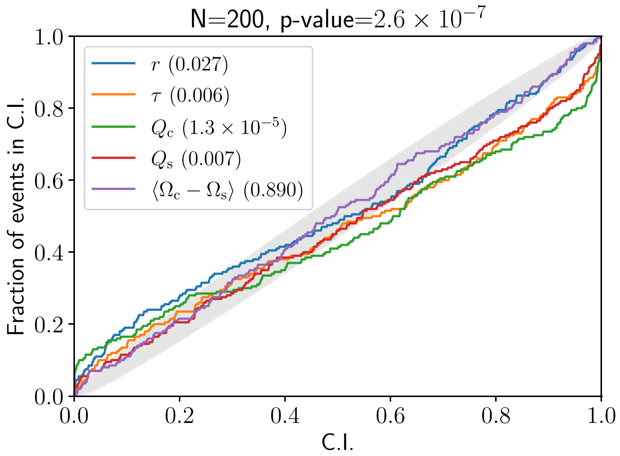

Figure F3. PP-plot analogous to Fig. D2 but for the 200 simulations in Fig. F2.

is safe, because the random noise contributes negligibly over a single rotation. We append t_2' to the sequence of TOAs and repeat up to t_3 and so on. To simulate TOA measurement errors, random Gaussian noise is added to the TOAs. We then take each triple of consecutive TOAs and their uncertainties (defined as the standard deviation of the added noise) and use TEMPO2 to fit a frequency and a frequency uncertainty.

The parameter estimation algorithm is applied to the fitted and exact Ω_c values generated by equations (F1)–(F3). We compare the two sets of results to assess the effect of the fitting process. As in Appendix D3, we sample the measurement error. In this section, the measurement uncertainties are different for each data point. We vary the measurement error by sampling over two parameters, α and β, which correspond to the parameters EFAC and EQUAD commonly used in pulsar timing programs (Hobbs et al. 2006; Lentati et al. 2014). Given values of α and β, we change the measurement error variances from R to R' according to the rule

$$R' = \alpha R + \beta. \tag{F4}$$

We simulate ideal, data-rich conditions where 1000 TOAs are spread over 1000 d and the TOA uncertainties are all 10^{-16} s. The recovered τ and Q_c values for 200 simulations are shown in Fig. F1 for TEMPO2 fitted frequencies (left panels) and exact frequencies (right panels). We find that Q_c and τ are estimated less accurately for fitted than for exact frequencies. Q_c is slightly underestimated for fitted frequencies; it is consistently just below the red diagonal. The τ values are also biased for the fitted frequencies, as can be seen from the many recovered τ values lying near the bottom of the prior range.

We also simulate data with the same TOA spacings and measurement uncertainties as the real data in Fig. 1 to quantify the effect of local fitting on parameter recovery under those conditions. The results are shown in Fig. F2 and Fig. F3. The plots of injected versus recovered parameters in Fig. F2 give an indication of how accurately the parameters are estimated. In the top-left panel, the recovered τ values have a large vertical spread. The Q_c values in the bottom-left panel are also uncertain, especially for smaller injected Q_c values. The PP-plot in Fig. F3 shows that the parameter curves deviate significantly from the diagonal. However, the estimates do not seem to have a significant bias in either direction. So while the parameter estimates in Section 5 are uncertain due to the fitting procedure, they have no significant systematic bias. The large spread of the recovered parameters is due in large part to the small number of frequency data points. More data would give a greater level of convergence to the true values as seen in Fig. F1.

APPENDIX G: FULL POSTERIOR DISTRIBUTION FOR PSR J1359−6038

In this appendix, for the sake of completeness, we present in Fig. G1 a visualization of the posterior distribution $p(\boldsymbol{\theta}|\boldsymbol{Y})$ for all six parameters $\theta = \{r, \tau, Q_c, Q_s, \langle \Omega_c - \Omega_s \rangle, \langle \dot{\Omega}_c \rangle\}$. The figure is formatted as a traditional corner plot: the panels with contours display $p(\boldsymbol{\theta}|\boldsymbol{Y})$ marginalized over four out of six parameters, and the panels with histograms display $p(\boldsymbol{\theta}|\boldsymbol{Y})$ marginalized over five out of six parameters. The distinctly peaked histograms in columns 2, 3, 4, and 6 are reproduced in Fig. 2. The flatter histograms in columns 1 and 5 correspond to parameters that cannot be inferred reliably from the data, as discussed in subsection 5.5.

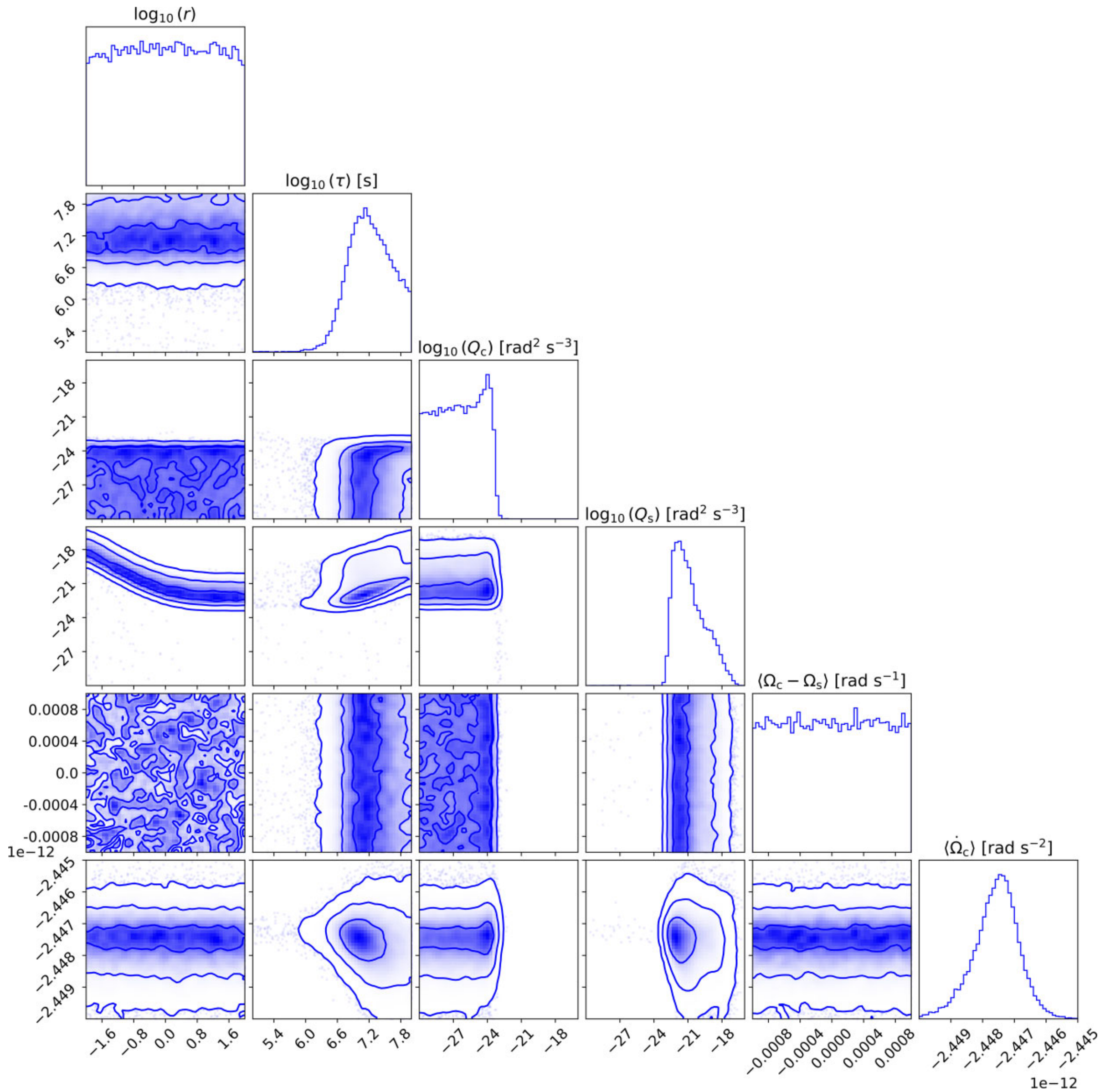

Figure G1. Six-dimensional joint posterior distribution for the two-component model for PSR J1359−6038. Marginalized posteriors for each parameter are shown as histograms on the diagonal. The lower half triangle of panels shows two-dimensional contour plots (with dark blue representing high probability) marginalized over all but two variables.

APPENDIX H: ONE-COMPONENT MODEL

In this appendix, Bayesian model selection is performed to determine whether the two-component model outperforms a simpler one-component model when modelling PSR J1359−6038.

The model given by (1) and (2) assumes there is a second component of the star, hidden from view, which couples to the crust, viz. the superfluid with angular velocity Ω_s. We could instead construct a simpler model where the pulsar is one rigid body described by the equation of motion

$$I\frac{d\Omega}{dt} = N + \xi(t), \tag{H1}$$

with

$$\langle\xi(t)\rangle = 0 \tag{H2}$$

$$\langle\xi(t)\xi(t')\rangle = \sigma^2\delta(t - t'), \tag{H3}$$

where Ω is the angular velocity of the crust (to which the pulses are tied), N is the constant spin-down torque, I is the pulsar's moment of inertia, and ξ is the stochastic torque responsible for timing noise modelled as a white noise process with strength σ.

We can make a Kalman filter for the one-component model and apply the same method of parameter estimation as in Section 3. For the one-component model, the update and measurement equations are still given by equations (6)–(9) with

$$F_i = 1, \tag{H4}$$

$$N_i = \frac{N}{I}\Delta t_i, \tag{H5}$$

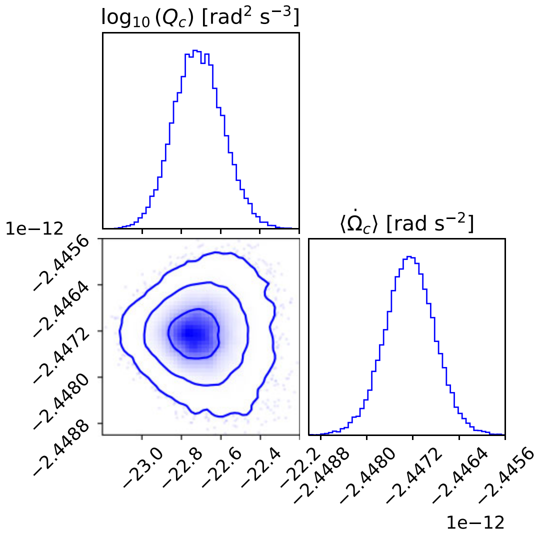

Figure H1. Two-dimensional joint posterior distribution for the one-component model for PSR J1359−6038. Marginalized posteriors for each parameter are shown as histograms on the diagonal. The bottom left panel shows the two-dimensional contour plot for the posterior.

$$Q_i = \frac{\sigma^2}{I^2}\Delta t_i, \tag{H6}$$

$$C = 1, \tag{H7}$$

and $\Delta t_i = t_i - t_{i-1}$.

We compare the two models' ability to explain the PSR J1359−6038 data by calculating the Bayesian evidence, Z, for each model given the data. We calculate the Bayesian evidence Z_M for a model M from the posterior distribution via the formula

$$Z_M = \int d\boldsymbol{\theta}\, p(\boldsymbol{Y}|\boldsymbol{\theta}, M) p(\boldsymbol{\theta}|M). \tag{H8}$$

The `dynesty` sampler generates samples from the posterior distribution to create the posterior plots and computes equation (H8) as a by-product. Let M_1 and M_2 denote the one- and two-component models, respectively. We assert no prior preference for either model so set $p(M_1) = p(M_2)$, yielding the evidence ratio (Bayes factor)

$$\frac{Z_1}{Z_2} = \frac{p(M_1|\boldsymbol{Y})}{p(M_2|\boldsymbol{Y})}. \tag{H9}$$

Equation (H9) is the relative preference for model 1 over model 2.

Fig. H1 shows the posterior distribution for the two parameters, $\langle\dot{\Omega}_\mathrm{c}\rangle = N/I$ and $Q_\mathrm{c} = \sigma_\mathrm{c}^2/I_\mathrm{c}^2$, obtained by applying the one-component model in equations (H1)–(H3) to the PSR J1359−6038 data from Section 4. The key results of the parameter recovery are summarized in Table H1. The contour plot in Fig. H1 shows no evidence for correlations between $\langle\dot{\Omega}_\mathrm{c}\rangle$ and Q_c.

The logarithms of the Bayesian evidences calculated for the one- and two- component models are $\log_{10}Z_1 = 449.89$ and $\log_{10}Z_2 = 456.70$, respectively. The relative log Bayes factor, $\log_{10}Z_2 - \log_{10}Z_1 = \Delta\log_{10}Z = 6.81$ categorically favours the two-component model. The uncertainties on $\log_{10}Z_1$ and $\log_{10}Z_2$ are $\sigma_1 = 0.017$ and $\sigma_2 = 0.017$, respectively, implying an approximate error for the log Bayes factor $\Delta\log_{10}Z$ of $(\sigma_1^2 + \sigma_2^2)^{1/2} = 0.02$.

Table H1. Static parameters inferred by the Kalman tracker and nested sampler for the one-component model in Appendix H, extracted from the two-dimensional joint posterior in Fig. H1. The fourth and fifth columns list two complementary measures of the dispersion in the posterior, viz. the full-width half-maximum (FWHM) and 90 per cent confidence intervals, respectively.

Parameter	Units	Peak value	FWHM interval	90 per cent confidence interval
$\log_{10}Q_\mathrm{c}$	rad^2 s^{-3}	-22.73	$(-22.86, -22.56)$	$(-22.92, -22.50)$
$\langle\dot{\Omega}_\mathrm{c}\rangle$	rad s^{-2}	-2.4473×10^{-12}	$(-2.4478 \times 10^{-12}, -2.4468 \times 10^{-12})$	$(-2.4480 \times 10^{-12}, -2.4465 \times 10^{-12})$

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