



# Noise analysis in the European Pulsar Timing Array data release 2 and its implications on the gravitational-wave background search

A. Chaluemau<sup>1,2,3★</sup>, S. Babak<sup>1,4★</sup>, A. Petiteau<sup>1,5</sup>, S. Chen<sup>1,2,3</sup>, A. Samajdar<sup>6</sup>, R. N. Caballero<sup>7</sup>,  
G. Theureau<sup>2,3,8</sup>, L. Guillemot<sup>2,3</sup>, G. Desvignes<sup>9,10</sup>, A. Parthasarathy<sup>9</sup>, K. Liu<sup>9</sup>, G. Shaifullah<sup>6,11</sup>,  
H. Hu<sup>9</sup>, E. van der Wateren<sup>12,13</sup>, J. Antoniadis<sup>9,14,15</sup>, A.-S. Bak Nielsen<sup>9,16</sup>, C. G. Bassa<sup>12</sup>,  
A. Berthereau<sup>2,3</sup>, M. Burgay<sup>17</sup>, D. J. Champion<sup>9</sup>, I. Cognard<sup>2,3</sup>, M. Falxa<sup>1</sup>, R. D. Ferdman<sup>18</sup>,  
P. C. C. Freire<sup>9</sup>, J. R. Gair<sup>19</sup>, E. Graikou<sup>9</sup>, Y. J. Guo<sup>9</sup>, J. Jang<sup>9</sup>, G. H. Janssen<sup>12,13</sup>, R. Karuppusamy<sup>9</sup>,  
M. J. Keith<sup>20</sup>, M. Kramer<sup>9,20</sup>, K. J. Lee<sup>7,9,21</sup>, X. J. Liu<sup>20,22</sup>, A. G. Lyne<sup>20</sup>, R. A. Main<sup>9</sup>,  
J. W. McKee<sup>23</sup>, M. B. Mickaliger<sup>20</sup>, B. B. P. Perera<sup>24</sup>, D. Perrodin<sup>17</sup>, N. K. Porayko<sup>9</sup>, A. Possenti<sup>17</sup>,  
S. A. Sanidas<sup>20</sup>, A. Sesana<sup>6,11</sup>, L. Speria<sup>25</sup>, B. W. Stappers<sup>20</sup>, C. Tiburzi<sup>12</sup>, A. Vecchio<sup>26</sup>,  
J. P. W. Verbiest<sup>9,16</sup>, J. Wang<sup>16</sup>, L. Wang<sup>21</sup> and H. Xu<sup>6,21,27</sup>

*Affiliations are listed at the end of the paper*

Accepted 2021 November 4. Received 2021 November 4; in original form 2021 October 19

## ABSTRACT

The European Pulsar Timing Array (EPTA) collaboration has recently released an extended data set for six pulsars (DR2) and reported evidence for a common red noise signal. Here we present a noise analysis for each of the six pulsars. We consider several types of noise: (i) radio frequency independent, ‘achromatic’, and time-correlated red noise; (ii) variations of dispersion measure and scattering; (iii) system and band noise; and (iv) deterministic signals (other than gravitational waves) that could be present in the PTA data. We perform Bayesian model selection to find the optimal combination of noise components for each pulsar. Using these custom models we revisit the presence of the common uncorrelated red noise signal previously reported in the EPTA DR2 and show that the data still supports it with a high statistical significance. Next, we confirm that there is no preference for or against the Hellings–Downs spatial correlations expected for the stochastic gravitational-wave background. The main conclusion of the EPTA DR2 paper remains unchanged despite a very significant change in the noise model of each pulsar. However, modelling the noise is essential for the robust detection of gravitational waves and its impact could be significant when analysing the next EPTA data release, which will include a larger number of pulsars and more precise measurements.

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

## 1 INTRODUCTION

Millisecond pulsars are remarkable for their long-term rotational stability, comparable with that of the most accurate atomic clocks on Earth (Matsakis, Taylor & Eubanks 1997; Verbiest et al. 2009; Hobbs et al. 2020). The highly beamed radio emission interacts with gravitational waves (GWs) causing deviations in the time of arrival of pulses observed by radio-telescopes. Pulsar timing arrays (PTAs) are used to search for ultralow-frequency (nHz–μHz) gravitational-waves by looking for their characteristic imprints on the times-of-arrival (ToAs) of the radio signals (Sazhin 1978; Detweiler 1979; Foster & Backer 1990; Perera et al. 2018).

The most promising GW source in the PTA frequency band is a population of nearby ( $z \leq 2$ ) slowly inspiralling supermassive ( $\geq 10^7 M_\odot$ ) black hole binaries (SMBHBs) with orbital periods  $\sim 0.1$ –10 yr. Those binaries were formed as a result of galaxy

collisions and their mergers will be observed in the LISA band, while PTA will see only the early and long lasting inspiral (Sesana et al. 2004). The most massive and closest binaries might be resolved as individual sources emitting continuous GWs (Sesana, Vecchio & Volonteri 2009; Babak et al. 2015), but the bulk of the SMBHB population emits rather weak GW signals which superpose and form a stochastic GW background (GWB) signal at low (nano-Hz) frequencies (Rajagopal & Romani 1995; Jaffe & Backer 2003; Wyithe & Loeb 2003). The key feature of this noise-like signal is a very specific spatial correlation in the data which depends only on the angular separation between two pulsars and is described by the Hellings–Downs curve (Hellings & Downs 1983) assuming general relativity.

Several PTA collaborations are working actively on the detection and characterization of such GWs. The three historical PTAs: Parkes pulsar timing array (PPTA; Kerr et al. 2020), the North American Nanohertz Observatory for Gravitational-waves (NANOGrav; Alam et al. 2021), and the European pulsar timing array (EPTA; Desvignes et al. 2016) cooperate with the emerging Indian pulsar timing array

\* E-mail: aurelien.chaluemau@cns-orleans.fr (AC); stas@apc.in2p3.fr (SB)

(InPTA)<sup>1</sup> as the International pulsar timing array (IPTA; Verbiest et al. 2016; Perera et al. 2019). The PPTA, NANOGrav, and the EPTA have recently published results produced using with their own independently processed data sets that show consistent evidence for the presence of a red (time-correlated) signal common among pulsars (Arzoumanian et al. 2020; Chen et al. 2021; Goncharov et al. 2021a) but without the necessary evidence for Hellings–Downs correlations, which would confirm the detection of a GWB. Encouraged by these promising results, the IPTA is planning to combine extended data sets from each PTA collaboration, in the hope that the GW nature of the observed signal can be confirmed (or disproved).

The biggest problem with PTA data is the lack of control of the noise: there are many potential sources which could contribute to the observed data which we model in a parametrized (and sometimes simplified) way when it is included in the ‘global fit’ (i.e. that fits all pulsars simultaneously). The high frequency end of PTA data is usually dominated by measurement white noise. The low frequency end is expected to be dominated by red noise processes. Besides the GWB discussed above, some of the red noise is also expected to be correlated among pulsars as a function of angular separation of each pair of pulsars on the sky (Tiburzi et al. 2015). Such sources of correlated noise are errors in the clock time standard (causing monopolar-type correlations) or systematic errors in the Solar system ephemeris (causing dipolar-type correlations). Moreover, we also expect the presence of spatially uncorrelated red noise, which is individual to each pulsar in the array. This is the spin noise (or timing noise), which is caused by the rotational variations of the pulsar arising from a variety of different phenomena (e.g. intrinsic processes, unmodelled objects in the vicinity of the neutron star, etc.). The red noise types described above are commonly referred to as *achromatic* red noise since they are independent of the observing radio frequency. Most of the PTA data also show the presence of *chromatic* red noise that depends on the radio frequency of the observations. In particular, we will consider the long-term variations of dispersion measure (DM), which add time delays to the ToAs as  $\Delta t \propto \nu^{-2}$ , and scattering variations ( $\Delta t \propto \nu^{-4}$ ), which are both caused by the time-varying electron column density between the pulsar and the radio telescope.

The sensitivity of PTA data to GWs is significantly affected by the levels of red noise and our ability to both detect and to characterize GW signals strongly depends on the faithfulness of the pulsar noise model (Caballero et al. 2016; Lentati et al. 2016; Hazboun et al. 2020; Goncharov et al. 2021b), which can vary significantly from pulsar to pulsar. Due to the large choice of possible noise components (we see it as various possible models of the noise) and their description or parametrization, the search for a GWB usually assumes a common (and simplified) noise model that is the same for each pulsar. The parameters of that simplified model are then inferred together with the parameters characterizing the GWB. It was shown (see for example Goncharov et al. 2021b, and the references therein) that the actual noise model could vary significantly from pulsar to pulsar, which could influence the detectability of a GWB (Hazboun et al. 2020). This is the main motivation behind this paper. We consider this paper as a companion of Chen et al. (2021), where the main focus was on finding and characterizing the common red signal (CRS) using two independent pipelines.

Based on previous exploratory investigations for each pulsar in the DR2, we suggest a finite set of noise models and use Bayes factors as a ranking statistic to choose between them, assuming that

all models are equally probable a priori. When the Bayes factor is not informative (close to one) we make the selection based either on the simplicity of the model or on the basis of computational efficiency with very few exceptions, which we explicitly discuss below. Once we had customized the noise model for each pulsar, we reproduced the main results of Chen et al. (2021), namely we confirmed the high statistical significance of a common red signal, but without sufficient evidence for a GWB induced Hellings–Downs correlation. We want to emphasize that the final noise model used for each pulsar was quite different from the simpler model assumed in Chen et al. (2021) for all pulsars. The noise model selection method described in this paper will be used on the extended (25 pulsars) EPTA data.

The rest of the paper is organized as follows. We give a brief description of DR2 data in the section 2. Section 3 gives a detailed description of each noise process that will be used in building the noise model. Section 4 summarizes the Bayesian framework used in this paper. The details of the single-pulsar noise model selection are given in Section 5. In Section 6 we consider the presence of a common red noise and we summarize our results in Section 7.

## 2 BRIEF DESCRIPTION OF EPTA DR2

The EPTA Data Release 2 (DR2) – six pulsars data set (Chen et al. 2021) – comprises up to 24 yr of high cadence observations of PSRs J0613–0200, J1012+5307, J1600–3053, J1713+0747, J1744–1134, and J1909–3744. These pulsars are observed in single-dish mode at four European radio telescopes: the Effelsberg 100-m radio telescope (EFF), the Nançay Radio telescope (NRT), the Lovell Telescope at the Jodrell Bank Observatory (JBO), and the Westerbork Synthesis Radio Telescope (WSRT). In addition, EPTA DR2 includes data from the Large European Array of Pulsars (LEAP), which is based on the combination of these four telescopes with the Sardinia Radio Telescope (SRT), forming a tied-array telescope (Bassa et al. 2015). The five radio telescopes contribute to all pulsars except PSR J1909–3744, which has a data set that contains only NRT observations because of its low -declination. Upgrades of telescopes, including improvements to or changes of receivers or backends, have been applied during the observational period, which makes the data set heterogeneous in timing precision and radio frequency coverage. We label the data by the telescope (or observatory) and the system that collected it, followed by the radio frequency in MHz (e.g. EFF.P200.1400). Having multiple systems in a PTA data set is both a curse, as we need to combine the data from all systems together taking into account possible systematics, and a blessing as the multiband observations are required to disentangle and characterize the chromatic noise and the system specific instrumental red noise (e.g. system noise; Lentati et al. 2016).

A characteristic ToA is computed from the time and frequency averaged profile of each observation, except for JBO.ROACH and NRT.NUPPI backends, which use, respectively, 2 and 4 (radio-frequency) sub-band ToAs per epoch. The ToAs of each pulsar are assembled together and used to fit the timing model (TM) parameters that describe the pulsars’s sky position and proper-motion, its spin frequency and corresponding derivative, and the DM and its two first derivatives. For pulsars in binary systems, the timing model also accounts for the orbital motion including Keplerian and post-Keplerian parameters. Phase jumps are included in the TM for each system and also for each of the JBO.ROACH & NRT.NUPPI.1484 sub-bands. The fits for the TM parameters were obtained using the TEMPO2 package (Hobbs, Edwards & Manchester 2006) with the JPL Solar system ephemeris DE438 [to transform the local observatory ToAs to the Solar system barycentre) and with the clock corrections

<sup>1</sup><https://inpta.gitlab.io/profile/index.html>

TT(BIPM2019) (time conversion from the observatory time standard to the Terrestrial Time (TT) given by the Bureau International des Poids et Mesures (BIPM)]. The end result are the *timing residuals*, i.e. the differences between the observed ToAs and those predicted by the TM, which are then analysed to search for GW signals.

### 3 MODELLING NOISE IN PTA DATA

In this section we describe the basic noise models. We will use an elegant description based on the Gaussian Process (GP) introduced in van Haasteren & Vallisneri (2014). Later, we will use those basic noise components to build a complete noise model for each pulsar in DR2 using Bayesian techniques.

Let us very briefly introduce the likelihood for Gaussian processes following van Haasteren & Vallisneri (2014). We assume that all noise components are Gaussian and stationary and we separate the white noise component  $\mathbf{N}$  from the rest. The Gaussian process can be introduced in two equivalent ways: (i) As a sum of deterministic basis functions  $\sum_i F_i(t)w_i$ , where  $w_i$  are weights – random Gaussian distributed variable  $\mathcal{N}(w_i^0, \Sigma_{ij})$ , where  $w_i^0$  is a mean value for each weight,  $\Sigma_{ij}$  is a covariance matrix, and  $F_i(t)$  are the basis functions. This is the weight-space view. (ii) As a continuous function such that the ensemble average is  $\mathbb{E}[f(t)] = m(t)$  and the covariance is  $\mathbb{E}[(f(t) - m(t))(f(t') - m(t'))] = C(t, t')$ . This is the function space view. Those two descriptions are related via

$$C(t, t') = \sum_{a,b} F_a(t) \Sigma_{ab} F_b(t'), \quad (1)$$

with  $a, b = 1, \dots, N$ . The red noise covariance matrix  $C(t, t')$  was introduced in van Haasteren & Levin (2012) and it was approximated using an incomplete Fourier basis in Lentati et al. (2013). Applying the Gaussian process approach to the PTA likelihood function we get (van Haasteren & Vallisneri 2014):

$$p(\delta t | w_a, \text{GP}) = \frac{e^{-\frac{1}{2} \sum_{ij} \delta t_i (-\sum_a F_a(t_i) w_a) (\mathbf{N}_{ij})^{-1} (\delta t_j - \sum_a F_a(t_j) w_a)}}{\sqrt{(2\pi)^n \det(\mathbf{N})}} \times \frac{e^{-\frac{1}{2} \sum_{a,b} w_a (\Sigma_{ab})^{-1} w_b}}{\sqrt{(2\pi)^m \det(\Sigma)}}, \quad (2)$$

where  $\delta t_i$  the  $i$ -th observed residuals with  $i, j = 1, \dots, n$ . The equivalent representation is given by

$$p(\delta t | \text{GP}) = \frac{e^{-\frac{1}{2} \sum_{ij} \delta t_i (\mathbf{N} + \mathbf{C}_{ij})^{-1} \delta t_j}}{\sqrt{(2\pi)^n \det(\mathbf{N} + \mathbf{C})}}, \quad (3)$$

where  $\mathbf{C}_{ij}^{rn} = \sum_{a,b} F_a(t_i) \Sigma_{ab} F_b(t_j)$ . The convenience of the latter description is that it can be computed efficiently using the Woodbury equality:

$$(\mathbf{N} + \mathbf{C})^{-1} \simeq (\mathbf{N} + \mathbf{F} \Sigma \mathbf{F}^T)^{-1} = \mathbf{N}^{-1} - \mathbf{N}^{-1} \mathbf{F} (\Sigma^{-1} + \mathbf{F}^T \mathbf{N}^{-1} \mathbf{F})^{-1} \mathbf{F}^T \mathbf{N}^{-1}. \quad (4)$$

In what follows we consider  $\mathbf{C}$  as a combination of several (chromatic and achromatic) red noise components each decomposed in its own set of basis functions.

#### 3.1 Marginalization over timing model

Before we introduce the noise components, we should explain how we treat the timing model. We assume that an initial fit of the timing model obtained with LIBSTEMPO (Vallisneri 2020) reduces it to a linear model where the coefficients are given by a design matrix. We analytically marginalize the likelihood over the TM parameter

errors described by that linear model. The analytic marginalization was first demonstrated in van Haasteren et al. (2009), but van Haasteren & Vallisneri (2014) have shown that it is equivalent to the marginalization of a corresponding Gaussian process with an improper prior.

The implementation of this marginalization in ENTERPRISE (Ellis et al. 2019; the package that we use throughout this project) uses the equivalence of weight space and function space description of a Gaussian process. The design matrices  $(M_a(t_i))$  are used as basis functions, the covariance for the TM process is given as  $\mathbf{C}^{\text{TM}} = \sum_{a,b} M_a(t_i) \Sigma_{ab}^{\text{TM}} M_b(t_j)$ , where the prior on the parameter errors is modelled as  $\Sigma = \lambda I$  with  $I$  being a diagonal unit matrix and  $\lambda$  is a large numerical number (see van Haasteren & Vallisneri 2014, for details). In the limit  $\lambda \rightarrow \infty$  this prior becomes improper, but in this analysis the values of  $\lambda$  are fixed but large so the prior is formally proper. The marginalization over timing errors ('weights') of equation (2) leads to equation (3) (this is a manifestation of the duality of the two descriptions).

The use of a very wide or improper prior in Bayesian model selection should be taken with great caution, especially when comparing two models where only one of them uses marginalization over the improper prior (this was also discussed in Chen et al. 2021). The penalization which is embedded in the prior (for being too wide) and propagates into the computation of the evidence is lost and reliable results from evidence-based model selection cannot be guaranteed. However in *all* noise models described below we perform marginalization over the TM parameters which brings them all to a common starting point for further comparison.

#### 3.2 White noise

As mentioned earlier, the white noise dominates the high frequency end of the PTA sensitivity band. The ToA errors ( $\sigma_{\text{ToA}}$ ) are estimated within the template-matching method (Taylor 1992) that is used to compute the ToAs. This method is based on the Fourier domain cross-correlation of a template profile with the integrated pulse profile at the corresponding epoch. The uncertainties of each ToA are further modified as

$$\sigma = \sqrt{E_f^2 \sigma_{\text{ToA}}^2 + E_q^2}.$$

EFAC ( $E_f$ ) is a multiplicative factor that takes into account ToA measurement errors (or radiometer noise). EQUAD ( $E_q$ ) is added in quadrature to account for any other white noise (such as stochastic profile variations Liu et al. 2012; Shannon et al. 2014; Lam et al. 2016) and for possible systematic errors. The white noise model therefore is given as

$$\mathbf{N} = (E_f^2 \sigma_{\text{ToA}}^2(t_i) + E_q^2) \delta_{i,j}, \quad (5)$$

where  $i$  and  $j$  indexing the ToAs of the corresponding backend. EFAC and EQUAD are phenomenological parameters that characterize the white noise for each system and for each pulsar.

#### 3.3 Stochastic red signals

It is essential for PTA analysis to properly describe the intrinsic red noise because of its possible correlation with low-frequency GW signals (Shannon & Cordes 2010). Results from simulations in Hazboun et al. (2020) have clearly demonstrated the impact of inaccurate red noise modelling on GWB results.

The single-pulsar stochastic red noise is a time-correlated signal modelled as a stationary Gaussian process. In this work we adopted

the ‘weight-space’ representation of the Gaussian Process. The timing residuals due to red noise at each epoch  $t_i$  are approximated as:

$$\delta t^{\text{SRS}}(t_i) = \sum_{l=1}^N X_l \cos(2\pi t_i f_l) + Y_l \sin(2\pi t_i f_l), \quad (6)$$

where  $X_l$  and  $Y_l$  are playing the role of weights and the basis functions are

$$\begin{aligned} F_{2l-1}(t_i) &= \cos(2\pi t_i f_l), \\ F_{2l}(t_i) &= \sin(2\pi t_i f_l), \end{aligned} \quad (7)$$

where  $l = 1, \dots, N$ . This representation would correspond to the usual Fourier transform if  $f_l = l/T$  (where  $T$  is the total time span) and we had regularly spaced epochs,  $t_i$ . However, the radio observations are quite irregular (besides maybe the last 5 yr or so) which makes the Fourier basis not exactly orthogonal. In addition, we do not use a complete set: we usually truncate it at some low frequency as we are interested in modelling the red noise. The optimal choice of frequencies was considered in van Haasteren & Vallisneri (2015); however, for all results presented here, we have used an evenly spaced  $\Delta f = 1/T$  set of frequencies, starting at  $f = 1/T$  and truncating at  $N/T$  where  $N$  is one of the parameters in the model selection.

The covariance matrix  $\Sigma$  for the Fourier coefficients (weights  $X_l$ ,  $Y_l$ ) is defined by the power spectral density (PSD),  $S(f)$ . The simplest model for the PSD of a stochastic red process in a single pulsar data set is a power law:

$$S_P(f; A, \gamma) = \frac{A^2}{12\pi^2} \left( \frac{f}{\text{yr}^{-1}} \right)^{-\gamma} \text{yr}^3, \quad (8)$$

where the amplitude  $A$  is the normalized value at the frequency of one over 1 yr ( $f = 1/\text{yr}$ ). The covariance matrix is given in the frequency domain by

$$\Sigma_{k\alpha, l\beta} = S_P(f_k; A_\alpha, \gamma_\alpha) \delta_{kl} \delta_{\alpha\beta} / T \quad (9)$$

where  $k, l = 1, \dots, N$ , and  $\alpha, \beta$  are pulsar indices – we consider spatially uncorrelated red noise – therefore, we have placed the Kronecker delta-function on the right-hand side.

An alternative description takes into account that the data is dominated by white noise at high frequencies, and that is captured in the broken power law (Arzoumanian et al. 2020):

$$S_{\text{BPL}}(f; A, \gamma, \delta, f_b, \kappa) = \frac{A^2}{12\pi^2} \left( \frac{f}{\text{yr}^{-1}} \right)^{-\gamma} \left( 1 + \left( \frac{f}{f_b} \right)^{1/\kappa} \right)^{\kappa(\gamma-\delta)} \text{yr}^3, \quad (10)$$

in which  $A$  is the amplitude at frequency  $f = \frac{1}{\text{yr}}$ ,  $f_b$  is the transition frequency,  $\gamma$  and  $\delta$  are, respectively, the slopes below and above  $f_b$ , and  $\kappa$  defines the smoothness of the transition. Jumping a bit ahead, we will perform Bayesian analysis of the data with priors on  $\kappa$  and  $f_b$  as uniform  $\mathcal{U}(0.01, 0.5)$  and log-uniform  $\log_{10}\mathcal{U}(10^{-10}, 10^{-6})$ , and the priors on  $A$  and  $\gamma$  are the same as for the simple power-laws models given in Table 1. The high-frequency spectral index  $\delta$  is fixed at 0.

A completely different approach is not to impose any particular spectral shape but rather estimate it from the data itself: this is the free-spectrum method (Lentati et al. 2016) in which

$$S_{\text{FS}}(f_i; \rho_i) = \rho_i^2 T, \quad (11)$$

where  $\rho_i$  is the spectral amplitude at frequency  $f_i = i/T$ , with  $i = 1, \dots, N$ , in units of residuals. This modelling is particularly useful for understanding the spectral content and for interpreting the results for the red noise models given above. The number of parameters in the free-spectrum approach is equal to the number of Fourier bins and

**Table 1.** Models and priors used for the single-pulsar model selection.

| Model (abbrev.)             | Parameters              | Priors (or fixed val.)                                     |
|-----------------------------|-------------------------|------------------------------------------------------------|
| White-noise (WN)            | EFAC                    | $\mathcal{U}(0.1, 5)$                                      |
|                             | EQUAD [s]               | $\log_{10}\mathcal{U}(10^{-9}, 10^{-5})$                   |
| Achromatic red-noise (RN)   | $A_{\text{RN}}$         | $\log_{10}\mathcal{U}(10^{-18}, 10^{-10})$                 |
|                             | $\gamma_{\text{RN}}$    | $\mathcal{U}(0, 7)$                                        |
| DM variations (DMv)         | $A_{\text{DM}}$         | $\log_{10}\mathcal{U}(10^{-18}, 10^{-10})$                 |
|                             | $\gamma_{\text{DM}}$    | $\mathcal{U}(0, 7)$                                        |
| Scattering variations (Sv)  | $A_{\text{Sv}}$         | $\log_{10}\mathcal{U}(10^{-18}, 10^{-10})$                 |
|                             | $\gamma_{\text{Sv}}$    | $\mathcal{U}(0, 7)$                                        |
| Free-chromatic noise (FCN)  | $A_{\text{FCN}}$        | $\log_{10}\mathcal{U}(10^{-18}, 10^{-10})$                 |
|                             | $\gamma_{\text{FCN}}$   | $\mathcal{U}(0, 7)$                                        |
|                             | $\chi_{\text{FCN}}$     | $\mathcal{U}(0, 7)$                                        |
| System-noise (SN or DMv-SN) | $A_{\text{SN}}$         | $\log_{10}\mathcal{U}(10^{-18}, 10^{-10})$                 |
|                             | $\gamma_{\text{SN}}$    | $\mathcal{U}(0, 7)$                                        |
|                             | $\chi_{\text{SN}}$      | 0 or 2                                                     |
| Band-noise (BN)             | $A_{\text{BN}}$         | $\log_{10}\mathcal{U}(10^{-18}, 10^{-10})$                 |
|                             | $\gamma_{\text{BN}}$    | $\mathcal{U}(0, 7)$                                        |
| DM events (E)               | $A_{\text{E}}$ [s]      | $\log_{10}\mathcal{U}(10^{-10}, 10^{-2})$                  |
|                             | $\tau_{\text{E}}$ [day] | $\log_{10}\mathcal{U}(1, 10^{2.5})$                        |
|                             | $t_0$ [MJD]             | $\mathcal{U}(54650, 54850)$ or $\mathcal{U}(57490, 57530)$ |
|                             | $\chi_{\text{E}}$       | 1, 2, 4 or $\mathcal{U}(0, 7)$                             |
| Annual chrom. (Y)           | $A_{\text{Y}}$ [s]      | $\log_{10}\mathcal{U}(10^{-10}, 10^{-2})$                  |
|                             | $\phi_{\text{Y}}$       | $\mathcal{U}(0, 2\pi)$                                     |
|                             | $\chi_{\text{Y}}$       | 2 or $\mathcal{U}(0, 7)$                                   |

it is therefore computationally more expensive. The priors used for each  $\rho_i$  will be log-uniform:  $\log_{10}\mathcal{U}(10^{-10}, 10^{-4})$ .

In the rest of this subsection, we give more details on each specific type of red noise that will be included in the total noise budget for each pulsar.

### 3.3.1 Achromatic red noise

The achromatic red-noise (which we henceforth denote as RN) is commonly used in single-pulsar noise models in order to characterize the long-term variability of the pulsar spin. Also referred to as ‘timing noise’ or ‘spin noise’, RN is a dominant feature in the ToAs of younger pulsars, and several physical processes have been suggested to explain it, such as magnetospheric variability (e.g. Lyne et al. 2010; Tsang & Gourgouliatos 2013) or interactions between the pulsar’s superfluid core and solid crust (Cordes & Shannon 2010). The origins of RN in MSPs may differ from that of young pulsars: due to their much weaker magnetic fields, superfluid turbulence has been suggested as a possible contributor to the RN in MSPs (Melatos & Link 2014).

We model RN using the descriptions given above: the power-law model will be our standard approach; however, we will also use the free spectrum and broken power law (red noise becomes white after some frequency) to guide the selection of the truncation frequency in the sum given by equation (6). This noise component is unique, independent of the observational radio frequency, and uncorrelated between different pulsars.



### 3.3.2 Chromatic red noise

During its propagation, the pulsar radio emission passes through and interacts with the ionized interstellar medium (IISM), the Solar system interplanetary medium and the Earth's ionosphere, which leads to frequency-dependent delays on the observed signal.

An important effect is the interstellar dispersion that induces a delay in the arrival time  $\Delta t^{\text{DM}} \propto \nu^{-2} \text{DM}$ , where  $\nu$  is the radio observing frequency and DM is the dispersion measure, which is the path integral of the free-electron density (Lorimer & Kramer 2004). This effect is taken into account during the observations and inside the timing model which considers its value at a reference epoch together with its first and second derivatives. However, the inhomogeneous and turbulent nature of the IISM also induces chromatic (i.e. dependent on the observing radio frequency  $\nu$ ) red noise that are important on the decade-long time-scales of PTA data (You et al. 2007; Keith et al. 2013). In addition, the orbital motion of the Earth around the Sun may induce an additional deterministic chromatic signal.

Another result of the radio signal's interaction with the IISM are scattering variations (Sv), corresponding to the multipath propagation of the radio signal due to diffraction and refraction in the IISM (Lorimer & Kramer 2004; Lyne et al. 2010). This causes a chromatic pulse broadening and a time delay with  $\sim \nu^{-4}$  chromaticity. The scattering variations are described as a stochastic red signal such that  $\Delta t^{\text{Sv}} \propto \nu^{-4}$ .

We describe phenomenologically any general chromatic red noise using the basis functions

$$\mathbf{F}_j^{\text{chrom.}}(t_i) = \mathbf{F}(t_i) * \left( \frac{\nu_j}{1.4 \text{ GHz}} \right)^{-\chi}, \quad (12)$$

where  $\mathbf{F}$  is the incomplete set of sin/cos basis functions,  $\nu_j$  is an observational radio frequency for a corresponding residual at the epoch  $t_i$ , and  $\chi$  is the chromatic index. We use the same covariance matrices for chromatic red noise as for achromatic red noise (power law, broken power law, free spectrum). It is essential to have multiband radio observations to disentangle chromatic from achromatic red noise (otherwise they are completely degenerate; see for example Caballero et al. 2016).

During model selection, we will consider the following chromatic red processes: (i) dispersion measurement variations (DMv) with  $\chi = 2$ ; (ii) scattering variations (Sv) with  $\chi = 4$ ; and (iii) a phenomenological 'free chromatic noise' model (FCN) with  $\chi$  taken to be a free parameter with prior  $\mathcal{U}(0, 7)$ . The FCN was first introduced in Goncharov et al. (2021b) and is used here as a diagnostic to verify the combined noise model.

### 3.3.3 System and band noise

The EPTA DR2 data set is a combination of ToAs produced by five radio telescopes which use different systems observing at radio frequencies ranging from  $\sim 300$  MHz to  $\sim 5$  GHz. Following Lentati et al. (2016) we introduce 'system' and 'band' red noise. The system noise (SN) term is a stochastic red signal specific to a single receiver system. Such a signal could, for example, arise from a miscalibration of polarizations or specific radio frequency interferences. We model this process as a stochastic red noise applied to the ToAs of only one considered system at a time. This noise is considered to be achromatic for every system except for NRT.NUPPI.1484 that is divided into four sub-bands and will be probed for the presence of both chromatic red process SN and DMv (labelled as DMv-SN).

The band noise (BN) is a stochastic red noise assigned to a specific radio frequency band. This is to account either for a possible frequency-dependent DM in the amplitude (additional to the overall

$\nu^{-2}$  factor) caused by multipath propagation of radio emission (Cordes, Shannon & Stinebring 2016) or by frequency-dependent calibration errors (van Straten 2013). Given the frequency coverage of the EPTA DR2 data set (Chen et al. 2021), we consider four radio bands for the BN:

- (i) Band.1:  $< 1$  GHz
- (ii) Band.2:  $[1, 2]$  GHz
- (iii) Band.3:  $[2, 3]$  GHz
- (iv) Band.4:  $> 3$  GHz.

### 3.4 Common red noise

In this subsection we describe the red noise common to all pulsars and differentiate between two groups of common stochastic red signals: (i) a common spatially uncorrelated red noise (CURN) signal; and (ii) a correlated common red noise signal. The CURN shares spectral properties across all pulsars but does not appear with any particular spatial correlation (random) for each pair of pulsars, its covariance matrix is described as

$$\Sigma_{k\alpha, l\beta} = S_P(f_k; A_{\text{CURN}}, \gamma_{\text{CURN}}) \delta_{kl} \delta_{\alpha\beta} / T, \quad (13)$$

where the amplitude and spectral index ( $A_{\text{CURN}}, \gamma_{\text{CURN}}$ ) are the same for all pulsars.

On the other hand, the stochastic GWB, and the clock and ephemerides errors are examples of truly spatially correlated red signals. We describe in detail the GWB, which is the only correlated red process considered in this paper (clock and ephemerides errors in the EPTA DR2 were investigated in Chen et al. 2021 and their presence was not supported by the data). The dimensionless characteristic strain spectrum of the GWB is given as a power law (Maggiore 2000; Jenet et al. 2006) with a reference frequency at  $1 \text{ yr}^{-1}$ :

$$h_c(f) = A_{\text{GWB}} \left( \frac{f}{\text{yr}^{-1}} \right)^{\alpha_{\text{GWB}}}, \quad (14)$$

with  $A_{\text{GWB}}$  and  $\alpha_{\text{GWB}}$ , respectively, the GWB strain amplitude and spectral index. The corresponding PSD  $S_P^{\text{GWB}}(f)$  can then be written as

$$\begin{aligned} S_P^{\text{GWB}}(f) &= \frac{1}{12\pi^2} \frac{1}{f^3} h_c^2(f) \\ &= \frac{A_{\text{GWB}}^2}{12\pi^2} \left( \frac{f}{\text{yr}^{-1}} \right)^{-\gamma_{\text{GWB}}} \text{yr}^3 \end{aligned} \quad (15)$$

with  $\gamma_{\text{GWB}} = 3 - 2\alpha_{\text{GWB}}$ .

The spectral slope of a GWB generated by a population of SMBHBs on circular and GW-driven orbits (Jaffe & Backer 2003; Chen, Sesana & Del Pozzo 2017) is expected to be  $\alpha_{\text{GWB}} = -2/3$ , or  $\gamma_{\text{GWB}} = 13/3$ . We use the same incomplete set of Fourier basis function as for the achromatic red noise described above but with a covariance matrix with spatial correlation coefficients  $\Gamma(\theta_{ab})$  corresponding to the Hellings–Downs curve (Lee, Jenet & Price 2008):

$$\Sigma_{k\alpha, l\beta} = S_P(f_k; A_{\text{GWB}}, \gamma_{\text{GWB}}) \delta_{kl} \Gamma(\theta_{\alpha\beta}) / T, \quad (16)$$

where  $\theta_{\alpha\beta}$  is the angular separation of a pair of pulsars. This model describes an isotropic component of the GWB.

Note that it is the spatial correlation that distinguishes a GWB from the CURN; therefore to clearly identify this it is necessary to infer  $\Gamma(\theta_{ab})$ . This is not possible with six pulsars and with the present EPTA timing precision as shown in Chen et al. (2021).

In addition, a CURN could be mimicked by the RN of the most sensitive pulsars, so it is crucial to study in detail the red processes in each pulsar in order to understand the significance of the CURN reported in Chen et al. (2021).

### 3.5 Deterministic signals

In addition to stochastic processes, we also consider two types of deterministic signals of non-GW nature. We have used prior information about the (possible) presence of these signals in the data of some pulsars.

#### Exponential dips

Several pulsars have displayed exponential timing-residual dips (E), in which there is a sudden frequency-dependent advance in the ToAs. It is relevant to our analysis that such events have been observed at least twice for PSR J1713+0747 over our observational time span, in 2008 ( $\sim$  MJD 54757) and in 2016 ( $\sim$  MJD 57510).

The first event, reported in Coles et al. (2015), Zhu et al. (2015), and Desvignes et al. (2016), is interpreted as a ‘DM event’, i.e. a drop in the electron column density along the line of sight producing a sudden reduction of DM that returns to the previously observed level exponentially over time.

The second event, reported in Lam et al. (2018), was accompanied by a pulse shape change and corresponding chromatic index lower than 2, so it is not compatible with a DM-related process. It was proposed instead (Goncharov et al. 2021b) that this event is related to processes in the pulsar’s magnetosphere.

We model the exponential dip delay at epoch  $t_i$  and radio frequency  $\nu_k$  as:

$$d^E(t_i, \nu_k; A_E, \tau, t_0, \chi_E) = \begin{cases} 0, & \text{if } t_i < t_0 \\ A_E \left( \frac{\nu_k}{1.4 \text{ GHz}} \right)^{-\chi_E} \exp \left( -\frac{t_i - t_0}{\tau} \right), & \text{if } t_i \geq t_0, \end{cases} \quad (17)$$

where  $A_E$  is the amplitude in residual units,  $t_0$  is the reference epoch of the event,  $\tau$  is the relaxation time, and  $\chi_E$  is the chromatic index, either fixed or being a free parameter with prior  $\mathcal{U}(0, 7)$ .

#### Annual chromatic signals

The second deterministic signal which could be present in the data is an annual chromatic process (which we label as ‘Y’) that results from electron density variations as the line of sight to the pulsar changes during the annual Earth motion around the Sun.

Previous investigations (Keith et al. 2013; Main et al. 2020) indicated such a signal is present in PSR J0613–0200, which we model as (Lentati et al. 2016; Goncharov et al. 2021b):

$$d^Y(t_i, f_i; A_Y, \phi, \chi_Y) = A_Y \left( \frac{f_i}{1.4 \text{ GHz}} \right)^{-\chi_Y} \sin \left( 2\pi \frac{t_i}{\text{yr}} + \phi \right), \quad (18)$$

where  $A_Y$  is the characteristic amplitude in residual units,  $\chi_Y$  is the chromatic index, and  $\phi$  is the initial phase. We consider either annual DM variations or annual scattering variations, with a chromatic index fixed at 2 or 4, respectively.

## 4 BAYESIAN INFERENCE FRAMEWORK

In this section, we briefly describe the Bayesian framework that will be applied to the model selection (see for example Sivia & Skilling 2006 for further reading). The main purpose of this section is to introduce notation that will be used in the following sections.

We will consider a set of models,  $\mathcal{M}_a$ , each characterized by parameters  $\theta_a$ , where the subscript  $a$  enumerates the models. The

probability of a given model  $\mathcal{M}_a$  given the observed residuals  $\delta t$  can be written using Bayes theorem:

$$P(\mathcal{M}_a | \delta t) = \frac{P(\delta t | \mathcal{M}_a) \pi_{\mathcal{M}_a}}{P(\delta t)}, \quad (19)$$

where  $\pi_{\mathcal{M}_a}$  is the prior probability of model  $\mathcal{M}_a$ ,  $P(\delta t | \mathcal{M}_a)$  is the probability of observing  $\delta t$  assuming that model  $\mathcal{M}_a$  is the correct one (this is the evidence of model  $\mathcal{M}_a$  and we denote it as  $\mathcal{Z}_{\mathcal{M}_a}$ ), and  $P(\delta t) = \sum_b P(\delta t | \mathcal{M}_b) \pi_{\mathcal{M}_b}$  is the probability of the observed data set, which we consider as a normalization factor. We have used previously published results as a guide for selecting models for a given pulsar, assuming that all considered models have equal prior probabilities unless otherwise specified. The model selection is based on the odds ratio:

$$\frac{P(\mathcal{M}_a | \delta t)}{P(\mathcal{M}_b | \delta t)} = \frac{\mathcal{Z}_{\mathcal{M}_a} \pi_{\mathcal{M}_a}}{\mathcal{Z}_{\mathcal{M}_b} \pi_{\mathcal{M}_b}} \quad (20)$$

Since we use equal priors, the odds ratio reduces to the *Bayes factor*  $\mathcal{B}_{\mathcal{M}_b}^{\mathcal{M}_a} = \mathcal{Z}_{\mathcal{M}_a} / \mathcal{Z}_{\mathcal{M}_b}$ .

For a given model the posterior on the parameters  $\theta_a$  is given again by Bayes theorem:

$$P(\theta_a | \delta t, \mathcal{M}_a) = \frac{P(\delta t | \theta_a, \mathcal{M}_a) P(\theta_a | \mathcal{M}_a)}{P(\delta t | \mathcal{M}_a)} = \frac{\mathcal{L}(\delta t | \theta_a, \mathcal{M}_a) \pi(\theta_a | \mathcal{M}_a)}{\mathcal{Z}_{\mathcal{M}_a}}, \quad (21)$$

where we will use the likelihood  $\mathcal{L}(\delta t | \theta_a, \mathcal{M}_a)$  marginalized over the timing model parameters (see discussion in the previous section) and  $\pi(\theta_a | \mathcal{M}_a)$  are priors on the model parameters. The evidence of a given model is computed from the fully marginalized posterior (Sivia & Skilling 2006):

$$\mathcal{Z}_{\mathcal{M}_a} = \int d\theta_a \mathcal{L}(\delta t | \theta_a, \mathcal{M}_a) \pi(\theta_a | \mathcal{M}_a). \quad (22)$$

Our decisions are based on the scale proposed in Jeffreys (1961), which is  $\mathcal{B}_{\mathcal{M}_b}^{\mathcal{M}_a} > 100$  indicates a preference for the model  $\mathcal{M}_a$  against  $\mathcal{M}_b$  with ‘decisive’ evidence. This interpretation criteria has been set phenomenologically, and revised in Kass & Raftery (1995), which suggests using the threshold value of 150 ( $\log_{10} \mathcal{B}_{\mathcal{M}_b}^{\mathcal{M}_a} \gtrsim 2.2$ ). Therefore, we use the range  $2 < |\log_{10} \mathcal{B}_{\mathcal{M}_b}^{\mathcal{M}_a}| < 2.2$  as a selection criteria. In the case of a non-conclusive Bayes factor we follow the Occam principle and select the model with the lowest prior volume (or computational cost).

The dimensionality of models that we consider varies from 16 to 75 parameters. To infer the parameter posterior for each model, we have used several numerical tools: (i) a parallel tempering Markov Chain Monte Carlo (MCMC) sampler PTMCMCSAMPLER (Ellis & van Haasteren 2017); and (ii) MC<sup>3</sup> (<https://gitlab.in2p3.fr/stas/samplermc>), both based on the Metropolis–Hastings algorithm (Metropolis et al. 1953; Hastings 1970). For computation of the evidence we have used the DYNesty (Speagle 2020) package based on the nested sampling algorithm (Skilling 2004; Skilling 2006). In addition, we have used the hyper-model method (proposed in Carlin & Chib 1995, extended in Hee et al. 2015 and applied to PTA in Taylor, van Haasteren & Sesana 2020) to obtain Bayes factors without evidence evaluation. In that approach models and their corresponding parameters are sampled using a hyper-parameter that switches between the models. We compare both approaches whenever possible. For multichain MCMC runs, we check convergence by computing the Gelman–Rubin ratio (Gelman & Rubin 1992; Brooks & Gelman 1998). Finally, we use the Jensen–Shannon divergence (JSD; Manning & Schütze 1999) to compare marginalized posteriors across models. This is the symmetric version of the Kullback–Leibler

divergence (Kullback 1959), which is bound to the range  $0 < \text{JSD}(A||B) < \ln(2) \sim 0.69$ , with zero corresponding to two identical probability distribution functions.

In Table 1, we summarize all noise models, their parameters, and the prior range used in this paper.

## 5 CUSTOMIZING THE NOISE MODEL FOR EACH PULSAR

In this section, we describe the selection of the optimized noise model for each pulsar. The selection is performed in several steps in a partially iterative way. We start with the base model used in Chen et al. (2021) that contains only achromatic red noise and DM variation components. Both these models depend on the number of Fourier frequency bins we use for the basis functions, or, in other words, on the high frequency cut-off that roughly corresponds to the transition from the red-noise to the white-noise dominated region. The analysis in Chen et al. (2021) used the 30 and 100 lowest Fourier frequencies ( $k/T$ , where  $k = 1...30(100)$ ) for RN and DMv, respectively, for all six pulsars, as a reasonable balanced choice given exploratory tests.

We start with these two sources of noise and apply Bayesian model selection to find the number of frequency bins for each pulsar. We use a simple power-law model for each red noise, and we use the broken power-law and free-spectrum models for guidance to minimize the set of model trials. Next, we include stochastic chromatic noise and deterministic signals to the noise budget, and, finally we test for the presence of system and/or band noise.

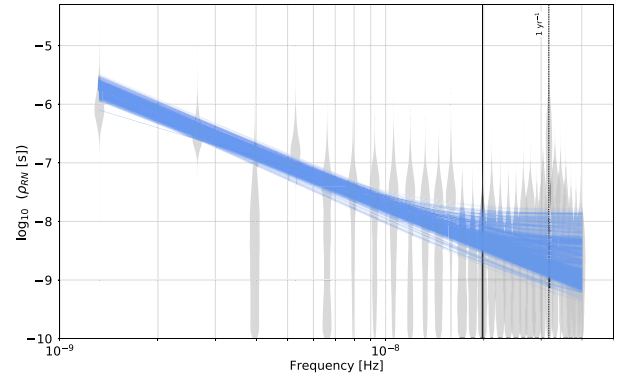
Each pulsar's noise model always includes white noise and we marginalize over the TM parameter errors (as implemented in ENTERPRISE).

### 5.1 Selection on number of Fourier modes for the achromatic red noise and DM variation

The importance of the choice of spectral binning has been discussed in van Haasteren & Vallisneri (2015), where the authors show limitations of the usual Fourier-sum approach with  $f_k = \{1/T, \dots, N/T\}$  in the presence of linear and/or quadratic signals, or if the stochastic red process spectral index is relatively high (i.e.  $\gamma \geq 7$ ). Earlier RN measurements have shown that such steeply rising noise is highly unlikely in MSPs and absent in the six pulsars considered here (see Verbiest et al. 2009; Caballero et al. 2016; Reardon et al. 2016). We therefore use a prior on the spectral index  $\mathcal{U}(0, 7)$ , and, as we will see later, this prior range is sufficiently broad. We have chosen to use Fourier frequencies  $f_k$  in our analysis.

We begin by identifying the most favoured number of Fourier bins for each pulsar. To do this, we assume that only the most commonly used noise components, RN and DMv, are present in each pulsar's data. This is the noise model that was assumed in Chen et al. (2021). We reassess the presence of those noise components in Section 5.2. We extend the short-hand notation for RN and DMv by appending the number of bins (basis functions) used in its description. For example, *RN30\_DMv100* refers to a model marginalized over the TM parameter errors, including white-noise parameters, and both RN and DMv with 30 and 100 Fourier modes, respectively.

The RN and DMv components could be highly correlated if we lack multiband observations. The RN dominates in the several lowest Fourier bins but is covered by the white noise at high frequencies. DMv, on other hand, depends on the observational radio frequency and therefore weakly (if at all) correlates with the white noise. Therefore, we should try to use a large number of Fourier bins to accommodate the dispersion information stored at high frequencies.



**Figure 1.** Achromatic red-noise spectrum of PSR J1713+0747 using the model *RN30\_DMv100*, with the free-spectrum PSD (grey violin plot) and broken power-law (blue) PSDs, showing 1000 random realizations of the posterior distribution. The vertical thin dotted line displays the  $1\text{yr}^{-1}$  frequency and the solid vertical line show the 15th bin.

For the RN, we expect that a relatively small number of bins contribute to the analysis before the white noise becomes dominant. For example, we have found that the use of *RN30\_DMv30* (which used to be the default choice) is disfavoured by a Bayes factor of more than  $10^5$  compared to the most favourable model *RN30\_DMv100* for PSR J1744–1134. In addition, the use of 30 bins for both noise components shows a very strong cross-model ‘leakage’ in the posterior of the RN and DMv parameters, which disappears completely in the favourable model.

We have analysed each pulsar using the broken power-law and free-spectrum models in order to get a rough indication on the expected range for the number of Fourier modes. Typical results of such an analysis (for RN) are given in Fig. 1 for PSR J1713+0747, where the estimation of the power in each free-spectrum bin is given by the grey violin-type histograms and we have overplotted 1000 realizations of the broken power law randomly drawn from the posterior as blue solid lines. The broken power-law model suggests that the transitional frequency  $f_b$  should be above 15 bins (as indicated by a vertical dashed line in Fig. 1). Therefore, for this pulsar we try 15, 20, and 30 Fourier modes for the RN. Similar analysis was performed for DMv and we decided to use 30, 50, 70, 100, and 150 modes (to choose from) for every pulsar, which allowed us to consider frequencies up to  $f_{\text{max}} \simeq 1/(2 \text{ months})$  for the pulsar with the longest data set (PSR J1713+0747).

We have performed Bayesian model selection across the pre-selected number of modes for both RN and DMv. The favourable models are summarized in Table 2. It shows that the data supports frequencies higher than  $100/T_{\text{span}}$  for DMv, except for J0613–0200 which displays no difference, either in evidence or in the posterior distribution of parameters, across the range of frequency-bin numbers for DMv that we have tried. As expected, for the RN, we require no more than 20 frequency bins for all pulsars except PSR J1012+5307, which requires 30. Where there was no clear preference between two (or more) models (Bayes factor less than 10), we preferred the model with the smallest number of bins for reasons of computational efficiency.

Let us give a few comments on the results presented in Table 2. Each quoted Bayes factor compares the selected model for the number of Fourier modes with the model used in Chen et al. (2021). One can see a significant gain in the Bayes factor for PSRs J1012+5307 and J1713+0747, which is mainly due to the extension of DMv to higher frequencies. In fact, all pulsars give



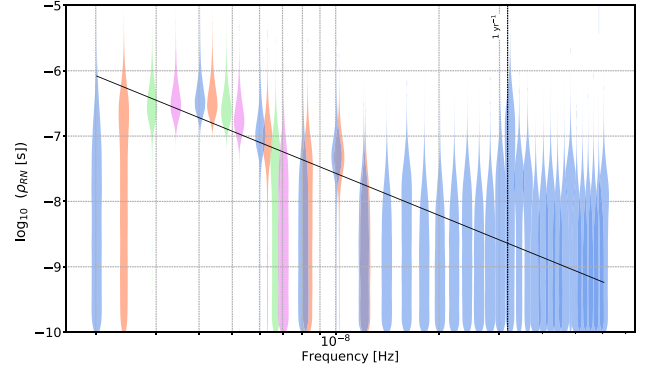
**Table 2.** Favoured number of Fourier modes for RN and DMv for the six pulsars. The third and fourth columns compare the favoured model with *RN30\_DMv100*, respectively, showing the Bayes factors and the Jensen–Shannon divergences for the RN amplitude (top) and spectral index (bottom). The models for PSR J1713+0747 also include the two exponential dips.

| Pulsar     | Favoured NBins     | $\log_{10} \mathcal{B}_{\text{RN30\_DMv100}}^{\text{Fav, NBins}}$ | J–S div. $A_{\text{RN}}$<br>$\gamma_{\text{RN}}$ |
|------------|--------------------|-------------------------------------------------------------------|--------------------------------------------------|
| J0613–0200 | <i>RN10_DMv30</i>  | 0.0                                                               | $2.31 \times 10^{-3}$<br>$1.32 \times 10^{-3}$   |
| J1012+5307 | <i>RN30_DMv150</i> | 3.1                                                               | $2.68 \times 10^{-2}$<br>$2.99 \times 10^{-2}$   |
| J1600–3053 | <i>RN20_DMv100</i> | –0.3                                                              | $4.44 \times 10^{-3}$<br>$4.03 \times 10^{-3}$   |
| J1713+0747 | <i>RN15_DMv150</i> | 6.3                                                               | $2.71 \times 10^{-3}$<br>$1.04 \times 10^{-3}$   |
| J1744–1134 | <i>RN10_DMv100</i> | –0.3                                                              | $4.85 \times 10^{-3}$<br>$4.66 \times 10^{-3}$   |
| J1909–3744 | <i>RN10_DMv100</i> | –0.1                                                              | $3.72 \times 10^{-3}$<br>$1.98 \times 10^{-3}$   |

a slight preference to 150 bins but with the Bayes factor close to one, and following our logic we have chosen to use 100 (30 for PSR J0613–0200) bins to reduce the computational cost. The negative, but close to zero, log-Bayes factor for PSRs J1600–3053, J1744–1134, J1909–3744 indicates that the use of 30 modes for RN has only a slight preference (what we would call *inconclusive*) and we have chosen the lowest allowed number of bins. The residuals in pulsar J0613–0200 data did not favour any particular number of modes. As we are mainly interested in the red noise, we have quantified the difference in the posterior of the red noise between two models by computing the Jensen–Shannon divergence. The last column indicates that there has been a statistically significant change in the RN relative to the base (Chen et al. 2021) models only for J1012+5307: the increase in the frequency range of the DMv process has constrained the RN to lower frequencies (the amplitude of the RN has slightly dropped while the spectral index has increased). We will revisit model selection for this pulsar in Section 5.2.

#### Red noise free-spectrum of PSR J1909–3744

In this paragraph we discuss PSR J1909–3744. This is one of the best timers, but it has the shortest observational span (about 11 yr) and the data was only acquired by NRT. The free spectrum and the power law (corresponding to the maximum a posteriori parameters) of the RN are plotted in Fig. 2. The blue violin plot shows the power distribution using our standard Fourier modes ( $i/T_{\text{span}}$ ). We performed additional runs (given by different colours) with scaled-down time span values as,  $T_{\text{span}} \rightarrow (T_{\text{span}}/1.2, T_{\text{span}}/1.45, T_{\text{span}}/1.7)$ , to get better resolution at low frequencies. Note that those frequency bins are not independent as we have used oversampling in the frequency domain, and the corresponding basis functions are not orthogonal even for evenly spaced data. One can clearly see that the spectrum flattens out and probably bends downwards at the lowest frequency bin. This bend is not very conclusive: the posterior at the lowest bin is poorly constrained, and could be caused by the gaps in the data (214 epochs from 2004 to 2011 with BON versus 695 from 2011 to 2020 with NUPPI backends). If this downturn is real, it could be related to processes intrinsic to the neutron star (Goncharov, Zhu & Thrane



**Figure 2.** Achromatic red-noise spectrum of J1909–3744 using the model *RN30\_DMv100*, with a power law (black solid line drawn with the maximum a posteriori), and four free-spectrum (violin plots) PSDs computed with four different minimum frequencies:  $i/T_{\text{span}}$  with  $i = 1$  (blue), 1.2 (orange), 1.45 (green), and 1.7 (pink). Here, 30 bins are drawn for the first one, 6 for the second, and 3 for the two others.

2020), or to a putative GWB produced by eccentric SMBHBs (Chen et al. 2017). We require a longer time span to better constrain the lowest frequencies and refine our interpretation.

## 5.2 Extending model selection to stochastic and deterministic signals

Let us now investigate the different red signals presented in Section 3. We start with the pulsar stochastic processes by (i) inspecting the presence of the chromatic signals in the data; and (ii) performing model selection to obtain the most favoured signal combination with RN, DMv, and Sv. The selection of the optimal number of modes for Sv was done in a manner similar to the one described in the previous section. We have also checked that the selected number of basis functions for RN and DMv is still optimal and found that to be the case for all pulsars except PSR J1600–3053, which we will discuss separately.

Next, we fit for the presence of an annual chromatic signal in PSR J0613–0200 data and exponential dips in PSR J1713+0747 (as discussed in Section 3.5). Finally, we search for the presence of system and band noise in each pulsar.

### 5.2.1 Stochastic chromatic signals

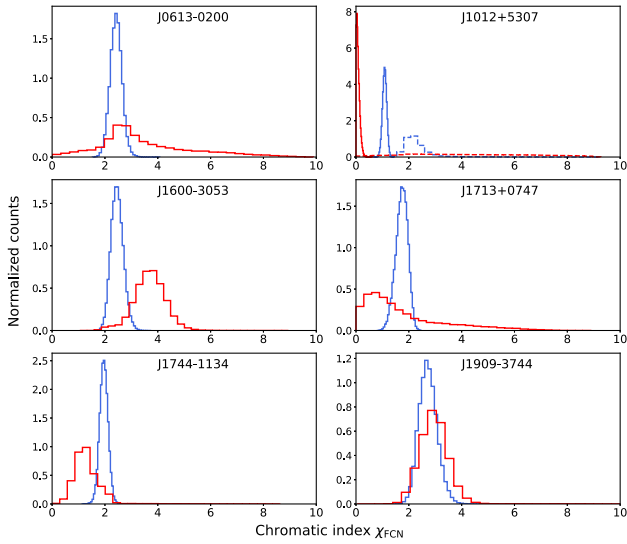
In the previous sub-section we assumed the presence of RN and DMv and concentrated on choosing the number of basis functions to describe the noise by a Gaussian Process. Now we fix the number of modes and check whether the data supports RN, DM, and Sv noise components. The probed models and the Bayes factors (with respect to the most favourable model indicated by the bold zeros) are summarized in Table 3. Below, we outline the procedure that we have followed.

In parallel to the direct computation of the evidence for each model, we have also conducted a noise diagnostic by using a noise model with RN and FCN (free chromatic index), which covers RN (chromatic index  $\chi_{\text{FCN}} = 0$ ), DMv ( $\chi_{\text{FCN}} = 2$ ), and Sv ( $\chi_{\text{FCN}} = 4$ ). The posterior on the chromatic index with *RN\_FC* is shown as blue histograms in Fig. 3 and in most cases it is centred around 2, indicating the presence of DMv, with two exceptions: PSR J1012+5307 (centred at 1.1) and J1909–3744 (centred at 3). We add DMv in our model and repeat the analysis with *RN\_DMv\_FC*. The FCN now



**Table 3.** Model selection for the stochastic chromatic signals. This table contains the log-10 Bayes factors of the highest evidence model ( $\log_{10} \mathcal{B} = 0.0$ ) over the set of other models that we have tried (given in columns). The selected model is indicated in bold. We have used 150 Fourier modes for RN in PSR J1012+5307 and 2 exponential dips are always included in the analysis of PSR J1713+0747.

| Pulsar     | <i>RN</i> | <i>DMv</i> | <i>Sv</i> | <i>RN_DMv</i> | <i>RN_Sv</i> | <i>DMv_Sv</i> | <i>RN_DMv_Sv</i> |
|------------|-----------|------------|-----------|---------------|--------------|---------------|------------------|
| J0613–0200 | –12.5     | –10.3      | –37.7     | <b>0.0</b>    | –2.2         | –1.7          | –0.3             |
| J1012+5307 | –25.0     | –63.0      | –143.7    | <b>0.0</b>    | –2.0         | –47.5         | 0.4              |
| J1600–3053 | –146.1    | –10.2      | –59.3     | –6.5          | –9.1         | <b>0.0</b>    | 0.0              |
| J1713+0747 | –36.8     | –42.6      | –125.0    | <b>0.0</b>    | –30.0        | –28.5         | –0.8             |
| J1744–1134 | –12.1     | –3.1       | –27.9     | <b>0.0</b>    | –10.7        | –2.4          | –1.8             |
| J1909–3744 | –66.7     | –82.1      | –244.4    | –2.1          | –3.4         | –21.0         | <b>0.0</b>       |



**Figure 3.** Marginalized posterior distributions of the chromatic index  $\chi_{\text{FCN}}$  in *RN\_FCN* (solid blue) and *RN\_DMv\_FCN* (solid red) for the six pulsars. The number of frequency bins for the RN and DMv power law are taken from Table 2. For PSR J1012+5307, we performed additional analyses with models *RN150\_FCN150* (dashed blue) and *RN150\_DMv150\_FCN150* (dashed red).

captures the remaining noise not covered by DMv and is indicated by the red histograms in Fig. 3. PSR J1600–3053 shows the presence of scattering noise ( $\chi_{\text{FCN}} = 4$ ). For this pulsar, variable and clumpy scintillation arcs in the secondary spectrum (i.e. the power-spectrum of the dynamic spectrum; see e.g. Cordes & Wolszczan 1986) are also seen in the *L*-band LEAP data (Main et al. in preparation), with power extending up to 16  $\mu\text{s}$  in delay, and averaged time delays at the 100ns level. These results add more confidence for the inclusion of this process, and indeed its presence is confirmed by the Bayes factor. However, including Sv into the noise model absorbs most of the red noise, making its presence inconclusive (as indicated in the table). PSRs J0613–0200, J1713+0747, and J1744–1134 show no sign of scattering variation noise. The chromatic index remains unchanged for J1909–3744, and the model selection indicates (though not very strongly) the presence of both DMv and Sv.

It is important to see how the inclusion of Sv changes the RN properties. In Fig. 4 we show the corner plot of the RN parameters for the model *RN\_DMv* in blue and the model *RN\_DMv\_Sv* in red, for the two pulsars that favour Sv. As mentioned above, the data is non-informative on the presence of RN in PSR J1600–3053 if we add Sv to the model and this can be seen by the poorly constrained posterior (red) in the left-hand panel. For J1909–3744, the inclusion

of Sv has a less drastic effect: it absorbs a small part of the RN at very low frequencies, which reduces the spectral index but pushes the amplitude slightly up.

### Peculiar red noise in PSR J1012+5307

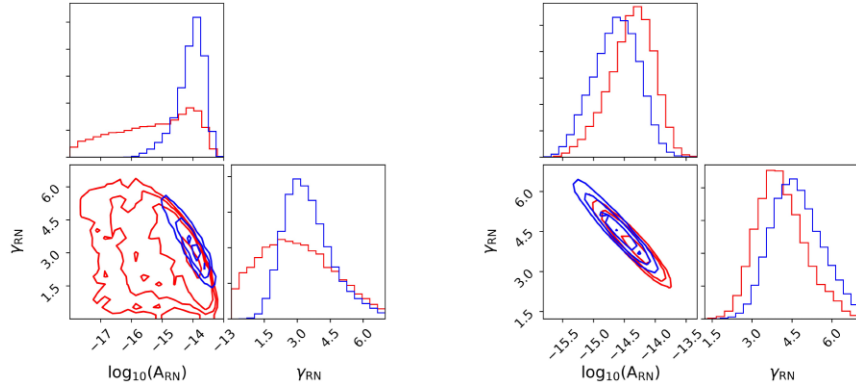
PSR J1012+5307 required a special investigation given the rather strong signal in model *RN\_DMv\_FCN* displayed in Fig. 3, with chromaticity index close to zero. It turned out to be unaccounted RN at high frequencies. In our initial analysis for the number of Fourier components we concentrated on low frequencies (up to 30 bins) for the RN and the FCN in *RN\_DMv\_FCN* picks up excess red noise which extends also to high frequencies.

Based on these findings we revisited the selection of Fourier modes done in Section 5.1 for this pulsar by allowing the RN to go up to 150 Fourier bins. The favoured model *RN150\_DMv30* has a Bayes factor  $\log_{10} \mathcal{B}_{\text{RN150_DMv30}}^{\text{RN150_DMv30}} = 29.8$  over the previous one, and therefore we adopt it in further investigations. In Fig. 5 we show the evolution of the red noise parameters (amplitude and spectral index) as we move from *RN30\_DMv150* (blue) to the new model *RN150\_DMv30* (red). The parameters of the RN process are better constrained, the red noise is significantly shallower (to accommodate the high frequency contribution) but the amplitude is slightly higher. The free spectrum estimation for this pulsar can be seen in the second panel of the right-hand column of Fig. A1. One can clearly see the low frequency red noise (well constrained power at the three lowest Fourier bins), but we also observe significant fluctuations at higher frequencies. The high frequency red noise is also evident in the time realization of this signal in the corresponding plot in the left-hand column. The red noise at high frequencies flattens out the power law of the overall RN process. However, we do not exclude the possibility that the red noise comprises two components of different origin.

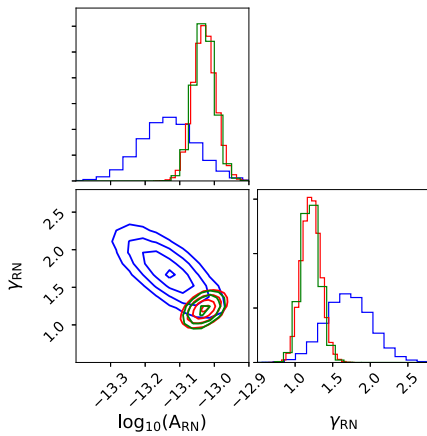
Using this number of modes, we have repeated the analysis of the data with the *RN150\_FCN150* and *RN150\_DMv30\_FCN150* models. The results are presented in Fig. 3 as dashed lines and confirm that *RN150\_DMv30* is sufficient to describe the data. The green histogram in Fig. 5 shows that adding Sv does not change the properties of the RN process.

### 5.2.2 Deterministic chromatic signals

The scattering and scintillation effects for J0613–0200 have been studied by Main et al. (2020), who also discuss the presence of annual variations of the arc curvatures. Furthermore, Keith et al. (2013) has shown the presence of annual chromatic signals which were reported in Goncharov et al. (2021b), suggesting the need for a noise model that includes an annual DM process (i.e.  $\chi_y = 2$ ). However, we



**Figure 4.** 2D distributions of achromatic red-noise amplitude and spectral index. Left plot: J1600–3053, the model  $RN\_DMv$  is in blue and  $RN\_DMv\_Sv$  is red; Right plot: J1909–3744, with  $RN\_DMv$  in blue and  $RN\_DMv\_Sv$  (favourable) in red.



**Figure 5.** 2D distribution of achromatic red-noise amplitude and spectral index for J1012+5307 with the noise models  $RN30\_DMv150$  (blue),  $RN150\_DMv30$  (red), and  $RN150\_DMv30\_SV150$  (green).

did not find conclusive evidence for the presence of such a signal, with Bayes factors  $\mathcal{B}_{RN\_DMv\_AnnualDM}^{RN\_DMv\_AnnualDM} = 8.3$  and  $\mathcal{B}_{RN\_DMv\_AnnualSv}^{RN\_DMv\_AnnualSv} = 7.0$ , which we do not consider sufficiently significant to justify its inclusion in the noise model. Note that (at least part of) the annual variations in the timing residuals might be absorbed by the TM parameter fit (through the pulsar sky position and proper motion).

As for the exponential dip events, we found their presence in PSR J1713+0747 data with high statistical confidence: the log-10 Bayes factors were 30.5, 13.9, and 46.8 favouring models that include, respectively, one single event at MJD 54757, one single event at MJD 57510, and both those events together. We found a chromatic index very consistent with scattering variations ( $\chi_{E1} = 4.07^{+1.77}_{-1.13}$ , with errors corresponding to the 68 per cent confidence interval) for the first event (left-hand panel of Fig. 6), and an index  $\chi_{E2} = 1.00^{+0.56}_{-0.49}$  for the second event. The index for the second event is consistent with Goncharov et al. (2021b), who reported a profile change for this event and proposed a cause linked with the pulsar’s magnetosphere instead of an IISM process. The posteriors for both events are shown in Fig. 6. Note that the posteriors of  $t_0$  for both epochs are sharply constrained between the two consecutive ToAs that surround the actual event dates and it is not railing against the prior range (see Table 1). For the rest of this work, we fix the chromatic indices of both events at 4 and 1 as discussed above.

### 5.2.3 System and band noise

The data set for each pulsar is made from the combination of ToAs produced with different receiver systems integrated in different radio telescopes. It might happen that one (or several) of these systems unexpectedly introduce extra noise. The idea of the system-noise model is to check this hypothesis.

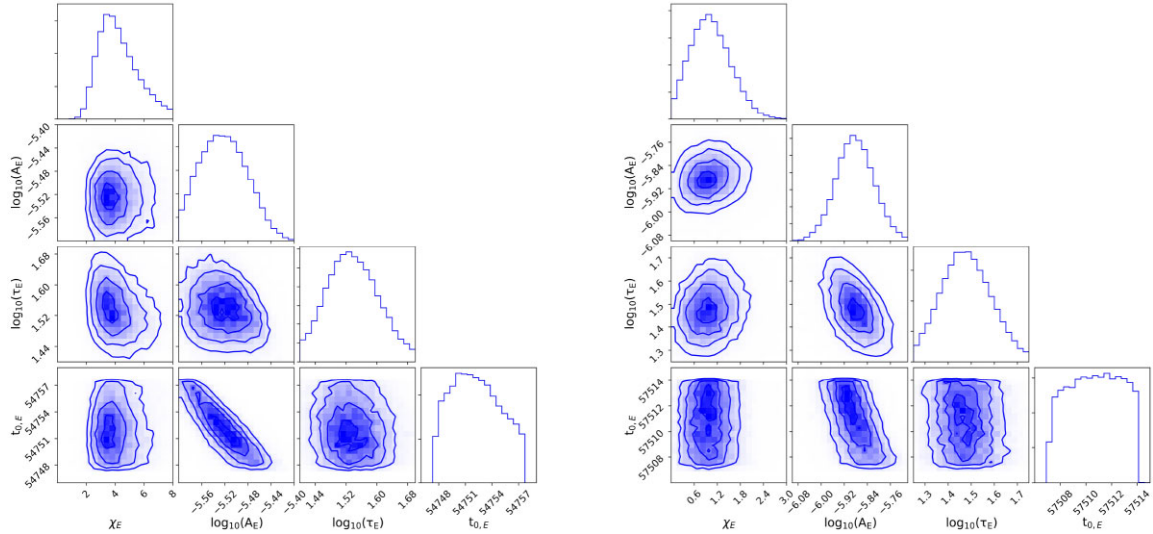
The EPTA data have a large number of systems, and checking all of them at the same time is computationally prohibitive. Instead, we have used an approach based on the hyper-model selection framework (Hee et al. 2015) to check for the presence of a noise in each system in turn. We introduce a switch (hyper) parameter which regulates the prior on the amplitude of the system noise being either significant/detectable or negligible. Those two models for each system imply that SN is always present but it could be at a detectable level or not. Note that the prior on the amplitude (log-uniform) is always taken into account and it has the same range for both models and therefore the main difference is in the likelihood, which gives some similarity to the dropout analysis (Arzoumanian et al. 2020). The posterior mean,  $P_i$ , of any given hyper-parameter indicates the probability of having red noise in the corresponding system. The ratio  $P_i/(1 - P_i)$  is the Bayes factor for the model including system noise in that component over the model without.

For this analysis, we exclude systems that have less than 3 yr of time span for any of the six pulsars. The excluded systems are JBO.DFB.1400, JBO.DFB.1520, WSRT.P1.323.C, WSRT.P1.367.C, WSRT.P1.840.C, WSRT.P1.1380.C, WSRT.P1.1380.1, WSRT.P1.1380.2.C, NRT.BON.1600, and NRT.NUPPI.1854. We also do not investigate NRT.NUPPI.2154, which contributes  $\sim 2.4$  yr for J1600–3053 and about 4.7 yr for J1909–3744, but with 16 epochs distributed in 1.1 yr, and only 2 epochs 4.6 yr after.

The data set of PSR J1909–3744 is composed by ToAs produced only from NRT observations, with three systems (BON backend) before MJD  $\sim 55812$  and four systems (NUPPI backend) after that date. This means that any possible system red noise will be totally correlated with RN, which will absorb it.

In Table 4 we report the log-10 of the Bayes factor in favour of including red noise in that component, computed as described above.

We observe three systems NRT.NUPPI.1484, JBO.ROACH.1520, and LEAP.1396 with a significant inclusion factor ( $10^5$ ,  $10^4$ , and  $10^4$ , respectively). We notice that these are three L-band systems active between 2011 and 2020, which corresponds to the major part of the data sets. The first two of the above mentioned systems (especially NRT.NUPPI.1484) are the largest contributors to the



**Figure 6.** Posterior distributions for the exponential dips found in J1713+0747 at MJD 54757 (left) and at MJD 57510 (right). The chromatic index  $\chi_E$  was used as a model parameter.

EPTA data and LEAP.1396 produces the ToAs with the lowest uncertainties (mean at  $1.86 \mu\text{s}$  for PSR J1012+5307 and lower than  $0.55 \mu\text{s}$  for the others). For PSR J1713+0747 two additional systems (EFF.S110.2639 and NRT.BON.2000) show signs of system noise, and it might also be present in NRT.NUPPI.2539 for PSRs J1012+5307 and J1744–1134.

The systems with an inclusion factor above 10 (presented in bold in Table 4) were selected for a detailed analysis of all possible combinations of system noise. We have found that DM-type chromatic system noise (DM-SN) is always favoured (in terms of Bayes factor) over the achromatic SN for the NRT.NUPPI.1484 system and it is included in the total noise budget for PSRs J0613–0200, J1012+5307, J1713+0747, and J1744–1134. It is not entirely clear why the data does not support its presence in J1600–3053 (where we have identified SN only in LEAP.1396). One plausible explanation is that the NRT.NUPPI.1484 ToAs dominate the data for this pulsar and to clearly disentangle SN from RN we would need to include data from other PTAs to test this assumption (that is a plan for the future IPTA data combination, repeating the previous such effort in Lentati et al. 2016).

Polarization calibration errors and radio frequency interference are possible causes for SN. Parameter posteriors for the SN of NRT.NUPPI.1484 and LEAP.1396 (Fig. 7) display overall consistency across pulsars, which corroborates the assumption of a red noise specific to these systems. We should emphasize again the presence of data from other systems (as expected in the IPTA data set) should greatly help to identify the system noise and disentangle it from the RN as was demonstrated in Lentati et al. (2016) for PSR J1730–2304 using the IPTA DR1 data set.

The results of the SN selection are presented in Table 5: inclusion of the SN leads to log-10 Bayes factors (15.9, 7.4, 13.8, 118.1, and 18.6) for PSRs J0613–0200, J1012+5307, J1600–3053, J1713+0747, and J1744–1134, respectively.

We switch now to the band noise investigation. The main radio frequency bands (cf. Table 4) in the data sets are Band.2 and Band.3, which contain the bulk of observations for all pulsars. Band.1 and Band.4 are only covered by one telescope: the WSRT and Effelsberg, respectively. As for SN, we do not investigate for BN if the corresponding time span of observations is less than 3

yr. Unfortunately, this made investigations in Band.1 and Band.4 inconclusive due to lack of sufficient data.

We found evidence of BN only in Band.3 of PSR J1713+0747 and Band.2 of PSR J1744–1134, with corresponding log-10 Bayes factors of 6.5 and 6.2. This result is somewhat consistent with Goncharov et al. (2021b), where Band-noise for both Band.2 and Band.3 were reported for these two pulsars. However, we have found that the inclusion of Band.2 for PSR J1713+0747 and Band.3 for PSR J1744–1134 results in insignificant log-10 Bayes factors, 0.6 in both cases. This points again to the importance of the IPTA data combination for the band noise analysis. Our results for BN are summarized in Table 5.

#### Poor constraint on RN for PSR J1744–1134

We revisited Table 3 after finding and fixing the set of noise sources included for each pulsar. We noticed that the RN for PSR J1744–1134 becomes poorly constrained using the  $RN\_DMv\_SN\_BN$  model (see red histograms in Fig. 8). After further investigation, we found that the Bayes factor  $B_{DMv\_SN\_BN}^{RN\_DMv\_SN\_BN} = 2$  hardly supports the presence of the RN. A similar result was found in Goncharov et al. (2021b), where the RN of this pulsar does not enter the favoured noise model. As another confirmation, the time-domain noise realizations of the red noise signal (see Fig. A1) are quite reduced for the  $RN\_DMv\_SN\_BN$  model (light grey) as compared to  $RN30\_DMv100$  (red). We have decided to keep the  $RN\_DMv\_SN\_BN$  model (as a conservative assumption); however, we will address its impact when we discuss the common red noise.

### 5.3 Summary of the model selection

The results of the noise model selection for each pulsar are summarized in Table 5. We report the RN parameters for each model, giving the median and 95 percent confidence interval. We also compare the custom-build noise model (M2) with the default noise model (M1) used in Chen et al. (2021). We quote the log-10 Bayes factors showing a strong preference for the custom noise model (ranging from 2 to 195). The biggest impact on the Bayes factor was caused by inclusion of the SN and BN components. Finally, we have evaluated

**Table 4.** Switch hyper-parameter (in log10-scale) for the achromatic system noise. Considered systems are given in the second column. The values above 1 (shown in bold), indicate evidence for the presence of red noise in that system. These systems were selected for a detailed analysis.

| Radio band | System                   | J0613–0200    | J1012+5307  | J1600–3053    | J1713+0747    | J1744–1134    | J1909–3744 |
|------------|--------------------------|---------------|-------------|---------------|---------------|---------------|------------|
| Band.1     | WSRT.P1.328              | −0.18         | −0.23       | –             | –             | –             | –          |
| <1 GHz     | WSRT.P1.328.C            | −0.23         | −0.07       | –             | –             | –             | –          |
|            | WSRT.P1.382              | −0.07         | −0.10       | –             | –             | –             | –          |
|            | WSRT.P1.382.C            | −0.02         | −0.01       | –             | –             | –             | –          |
|            | WSRT.P1.840              | –             | –           | –             | −0.15         | –             | –          |
|            | WSRT.P1.840.C            | –             | –           | –             | −0.07         | –             | –          |
|            | WSRT.P2.350              | –             | −0.11       | –             | −0.16         | –             | –          |
| Band.2     | EFF.EBPP.1360            | −0.16         | −0.15       | –             | 0.24          | −0.23         | –          |
| [1,2] GHz  | EFF.EBPP.1410            | 0.41          | −0.16       | –             | −0.18         | −0.36         | –          |
|            | EFF.P200.1380            | –             | 0.54        | –             | –             | –             | –          |
|            | EFF.P200.1400            | −0.08         | –           | −0.28         | –             | –             | –          |
|            | EFF.P200.1400.np         | –             | –           | –             | −0.44         | –             | –          |
|            | EFF.P217.1380            | –             | 0.26        | –             | –             | –             | –          |
|            | EFF.P217.1400            | −0.09         | –           | −0.29         | 0.32          | −0.14         | –          |
|            | EFF.P217.1400.np         | –             | –           | –             | −0.09         | –             | –          |
|            | JBO.ROACH.1520 (JBO.1.5) | <b>1.93</b>   | <b>2.48</b> | 0.01          | <b>≥ 5.00</b> | <b>≥ 5.00</b> | –          |
|            | LEAP.1396 (LEAP.1.4)     | <b>≥ 5.00</b> | 0.42        | <b>≥ 5.00</b> | <b>≥ 5.00</b> | <b>≥ 5.00</b> | –          |
|            | NRT.BON.1400             | 0.66          | −0.23       | −0.27         | −0.13         | 0.62          | −0.33      |
|            | NRT.NUPPI.1484 (NUP.1.4) | <b>≥ 5.00</b> | <b>2.65</b> | <b>3.80</b>   | <b>≥ 5.00</b> | <b>≥ 5.00</b> | −0.41      |
|            | WSRT.P1.1380             | −0.21         | –           | –             | –             | –             | –          |
|            | WSRT.P1.1380.2           | –             | −0.18       | –             | −0.22         | –             | –          |
|            | WSRT.P2.1380             | −0.01         | −0.13       | −0.28         | 0.56          | −0.16         | –          |
| Band.3     | EFF.EBPP.2639            | 0.05          | −0.29       | –             | −0.20         | −0.19         | –          |
| [2,3] GHz  | EFF.S110.2487            | –             | −0.30       | –             | –             | −0.17         | –          |
|            | EFF.S110.2639            | 0.08          | –           | −0.25         | <b>2.63</b>   | –             | –          |
|            | NRT.BON.2000 (BON.2.0)   | −0.24         | −0.23       | −0.01         | <b>≥ 5.00</b> | −0.31         | −0.30      |
|            | NRT.NUPPI.2539 (NUP.2.5) | −0.23         | <b>1.53</b> | −0.34         | −0.24         | <b>1.48</b>   | −0.40      |
|            | WSRT.P1.2273.C           | –             | –           | –             | −0.10         | –             | –          |
|            | WSRT.P2.2273             | –             | –           | –             | −0.03         | –             | –          |
| Band.4     | EFF.S60.4850             | –             | –           | –             | −0.06         | –             | –          |
| >3 GHz     | EFF.S60.4857             | –             | −0.16       | –             | –             | 0.13          | –          |

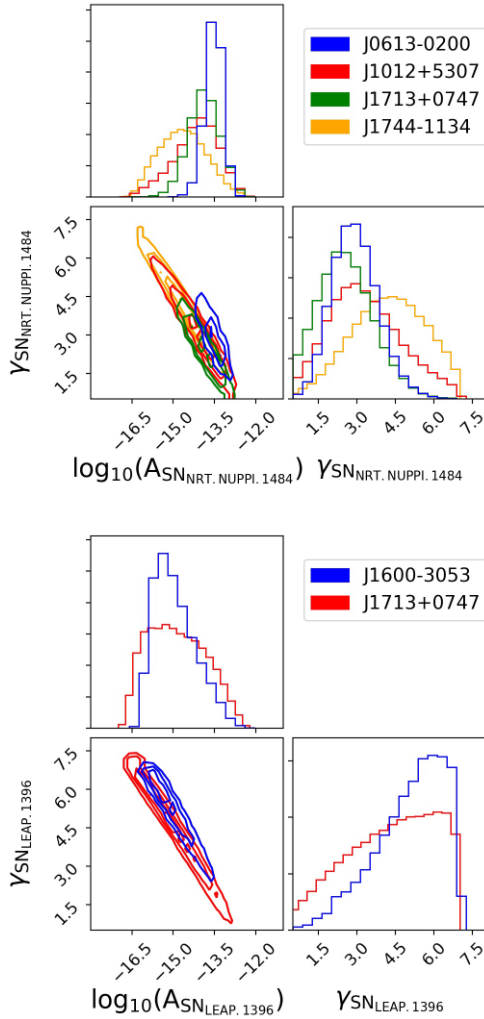
the Anderson–Darling diagnostic  $A^2$  (Anderson & Darling 1952) for the whitened residuals, obtained after subtracting the time domain noise realization drawn from the maximum likelihood from the residuals and dividing by the ToA uncertainties. While the Bayes factor measures the relative closeness of the whitened residuals to a Gaussian distribution, it does not tell us if the final result is actually Gaussian (in other words, if the favoured model is actually good). The Anderson–Darling statistic addresses precisely this question.  $A^2 = 2.5$  is a value at which one fails to reject the null hypothesis (i.e. following a Gaussian distribution) at 95 per cent confidence level. A lower value of the statistic corresponds to a better agreement with a Gaussian distribution. We see the overall improvement for the M2 model. The high values for PSR J1713+0747 could be caused by a few outliers in the whitened residuals. To test this assumption, we recompute the statistic after removing outliers found with the Grubbs’ test (Grubbs 1950), obtaining 3.2 and 2.1, respectively, for M1 (with 9 outliers) and M2 (with 6 outliers). The Anderson–Darling statistic for PSR J1744–1134 reduces to 1.0 if we exclude RN from the favourable (M2) model.

In preparation for the next section where we consider a common red signal, we investigated how much the white noise parameters impact the measurement of the RN. Following Lentati et al. (2015), we have performed a noise analysis with all white noise parameters fixed to the maximum-likelihood values and compared the RN posteriors to the previously obtained results. The results are quoted as Jensen–Shannon divergences (last column of Table 5) and show a very good consistency ( $J-S\,div < 3 \times 10^{-3}$ ). This confirmed that we can safely fix the white noise parameters for further investigations.

## 6 IMPACT ON THE SEARCH FOR A COMMON RED NOISE

In this section, we investigate how the custom single-pulsar noise model affects the results of the common red signal analyses reported in Chen et al. (2021). We consider here the CRS either with Hellings–Downs spatial correlations (GWB; Hellings & Downs 1983) or without (CURN).





**Figure 7.** 2D distributions of amplitude and spectral slope of NRT.NUPPI.1484 DM-SN (top) and LEAP.1396 SN (bottom), using the final favoured model for each pulsar.

As in the previous section, we denote the default base noise model used for all pulsars in Chen et al. (2021),  $RN30\_DMv100$ , as M1, and label the custom models (summarized in Table 5) for each pulsar as M2, and fix the white-noise parameters to their maximum-likelihood values (in the corresponding models). We model the common red noise using 30 uniformly spaced Fourier modes  $f_k = k/T_{\text{tot}}$ ,  $k = 1 \dots 30$ , where  $T_{\text{tot}}$  is the time span between the lowest and highest epoch from the combined data of all pulsars.

### 6.1 Contribution of each pulsar to the common red signal

Following Arzoumanian et al. (2020), Goncharov et al. (2021a), and Chen et al. (2021), we study the contribution of each pulsar data set to the inferred presence of a common red process using dropout analysis. One additional parameter is added to the model for each pulsar, with a uniform prior, and these are sampled as part of the model. When the parameter is one, the common red signal is included in the model for that pulsar used in the likelihood and when it is zero it is not. The dropout factor is the ratio of the fraction of posterior samples when the CRS is included to the fraction when it is not included.

The resulting dropout factors (see Fig. 9, blue hollow circles) with M1+CURN are very similar to those presented in fig. 5 in Chen et al. (2021). The dropout factor for PSR J1012+5307 is around 1 (consistent with the results of Arzoumanian et al. 2020) and most likely caused by abnormal red noise at high frequency (see discussion in 5.2.1).

The contribution of each pulsar to the common red noise has decreased for the custom model M2+CURN with the biggest drop shown for PSR J1600–3053 (the one which did not support RN). Despite that the overall result remains: these pulsars support the presence of a CURN. Interestingly, if we discard RN from the M2+CURN model of PSR J1744–1134 (see discussion in Section 5.2.3), it is picked up by the CURN leading to an increase in the drop-out factor (see hollow red circle in Fig. 9). Note that this poorly constrained (and poorly understood) signal could potentially affect the sensitivity to the GWB. Our choice to keep RN inside the M2 model for PSR J1744–1134 was a conservative choice.

### 6.2 Spectral properties of common red signal

Fig. 10 and Table 6 summarize the spectral properties of the CURN and GWB. We do not see significant changes in the new results (red) from those found previously (blue; Chen et al. 2021). The amplitudes of both CURN and GWB are slightly lower and the spectral indices are a bit shallower with the M2 models. The median amplitude is reduced from  $A_{\text{CURN},M1} = 5.42^{+4.48}_{-2.81} \times 10^{-15}$  to  $A_{\text{CURN},M2} = 4.88^{+4.94}_{-2.85} \times 10^{-15}$  (95 per cent credible interval) and the uncertainties of the spectral index are somewhat larger. The shift in the amplitude of the CURN is most likely due to partial absorption of the noise components which are unmodelled in M1, and which are accounted for in the extended set of M2. Similarly for the GWB we have  $A_{\text{GWB},M1} = 5.01^{+4.34}_{-2.63} \times 10^{-15}$  and  $A_{\text{GWB},M2} = 4.87^{+5.26}_{-2.88} \times 10^{-15}$ . Note that the amplitude of the CRS in M2 is the same for the CURN and GWB models.

We observe no changes in the CURN amplitude and spectral index posteriors using M2 with or without RN for PSR J1744–1134, with corresponding Jensen–Shannon divergences  $2.95 \times 10^{-3}$  and  $1.84 \times 10^{-3}$ .

We have also considered a free-spectrum model for CRS and plot it in Fig. 11. The M1+CRS model is presented by the hollow violins and we plot M2+CRS by filled blue violins. The left plot shows the addition of CURN to the M1 and M2 noise models. We notice the slight drop in the amplitude at the lowest frequency and slightly better constraint of amplitude at the second frequency bin. The free spectrum for GWB (right plot) shows a slight drop in power at the lowest frequency. These results nicely confirm our main findings with the power-law model of a decreased amplitude and shallower spectral index with M2 compared to M1.

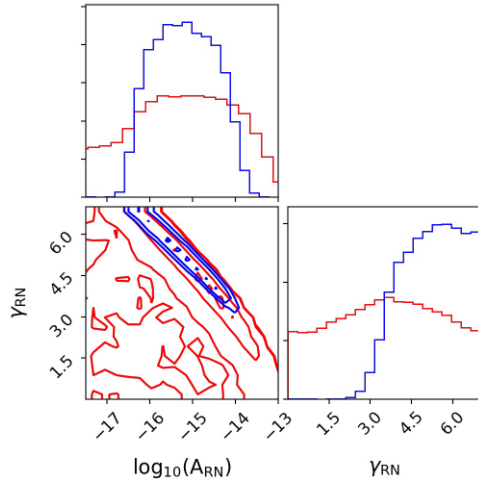
### 6.3 Effects on the statistical significance of the common red signal

Finally, we evaluate Bayes factors considering M1/M2+CURN/GWB against pulsar noise models without any common process (PSRN), M1/M2.

The results are summarized in Table 7. We have used two methods to compute Bayes factors (Dynesty, indicated by ‘Dyn.’ and the hyper-model, presented in Section 4, indicated by ‘Hyp’) for cross checking. First of all we have re-derived the results of Chen et al. (2021) for the M1 model: we observe strong evidence for the presence of a CRS but the data is not informative about its nature (could equally

**Table 5.** Final noise models for the six pulsars in EPTA DR2. The third and fourth columns show the median of the RN amplitude ( $A_{\text{RN}}$ ) and spectral slope ( $\gamma_{\text{RN}}$ ), with corresponding 95 per cent confidence interval. The fifth column displays the  $\log_{10}$  Bayes factors of the custom model over the model  $RN30\_DMv100$ . The sixth and seventh columns show the Anderson–Darling statistic applied to the whitened residuals for (i)  $RN30\_DMv100$ ; and (ii) the selected noise model. The last column shows the Jensen–Shannon divergences of the RN amplitude and spectral index posteriors of the selected noise model relative to the same model with fixed white-noise parameters.

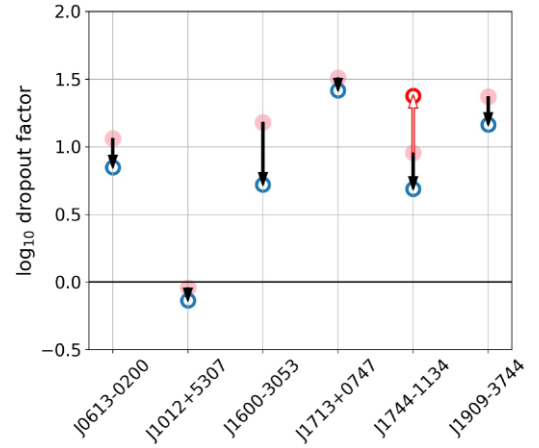
| Pulsar     | Sel. model               | $A_{\text{RN}}$          | $\gamma_{\text{RN}}$   | $\log_{10} \mathcal{B}_{\text{M1}}^{\text{M2}}$ | $A_{\text{M1}}^2$ | $A_{\text{M2}}^2$ | J–S $\text{div}_{\text{WNf}} A_{\text{RN}}$<br>$\gamma_{\text{RN}}$ |
|------------|--------------------------|--------------------------|------------------------|-------------------------------------------------|-------------------|-------------------|---------------------------------------------------------------------|
| J0613–0200 | <i>RN10 DMv30</i>        | $-14.93^{+1.26}_{-1.17}$ | $5.07^{+1.83}_{-2.34}$ | 15.9                                            | 0.4               | 0.3               | $2.36 \times 10^{-3}$                                               |
|            | <i>DMv-SN_NUP_1.4</i>    |                          |                        |                                                 |                   |                   | $1.82 \times 10^{-3}$                                               |
| J1012+5307 | <i>RN150 DMv30</i>       | $-13.03^{+0.08}_{-0.08}$ | $1.16^{+0.32}_{-0.29}$ | 40.7                                            | 1.8               | 1.4               | $6.60 \times 10^{-4}$                                               |
|            | <i>DMv-SN_NUP_1.4</i>    |                          |                        |                                                 |                   |                   | $3.80 \times 10^{-4}$                                               |
|            | <i>SN_NUP_2.5</i>        |                          |                        |                                                 |                   |                   |                                                                     |
| J1600–3053 | <i>DMv30 Sv150</i>       | –                        | –                      | 20.0                                            | 0.3               | 0.2               | –                                                                   |
|            | <i>SN_LEAP_1.4</i>       |                          |                        |                                                 |                   |                   | –                                                                   |
| J1713+0747 | <i>RN15 DMv150</i>       | $-14.50^{+0.51}_{-0.86}$ | $3.94^{+1.82}_{-1.13}$ | 195.4                                           | 5.5               | 4.1               | $1.76 \times 10^{-3}$                                               |
|            | <i>2 Exp. dips</i>       |                          |                        |                                                 |                   |                   | $1.92 \times 10^{-3}$                                               |
|            | <i>DMv-SN_NUP_1.4</i>    |                          |                        |                                                 |                   |                   |                                                                     |
|            | <i>SN_JBO_1.5</i>        |                          |                        |                                                 |                   |                   |                                                                     |
|            | <i>SN_LEAP_1.4</i>       |                          |                        |                                                 |                   |                   |                                                                     |
|            | <i>SN_BON_2.0</i>        |                          |                        |                                                 |                   |                   |                                                                     |
| J1744–1134 | <i>BN_Band.3</i>         |                          |                        |                                                 |                   |                   |                                                                     |
|            | <i>RN10 DMv100</i>       | $-15.31^{+2.03}_{-2.50}$ | $3.68^{+3.13}_{-3.46}$ | 22.6                                            | 1.0               | 1.2               | $4.39 \times 10^{-3}$                                               |
|            | <i>DMv-SN_NUP_1.4</i>    |                          |                        |                                                 |                   |                   | $1.79 \times 10^{-3}$                                               |
| J1909–3744 | <i>BN_Band.2</i>         |                          |                        |                                                 |                   |                   |                                                                     |
|            | <i>RN10 DMv100 Sv150</i> | $-14.45^{+0.66}_{-0.85}$ | $4.22^{+2.16}_{-1.65}$ | 2.1                                             | 0.8               | 0.6               | $3.10 \times 10^{-4}$<br>$3.20 \times 10^{-4}$                      |



**Figure 8.** 2D posterior distribution of the red-noise parameters for PSR J1744–1134 for model  $RN\_DMv$  (blue) and the most favoured model,  $RN\_DMv\_SN\_BN$  (red).

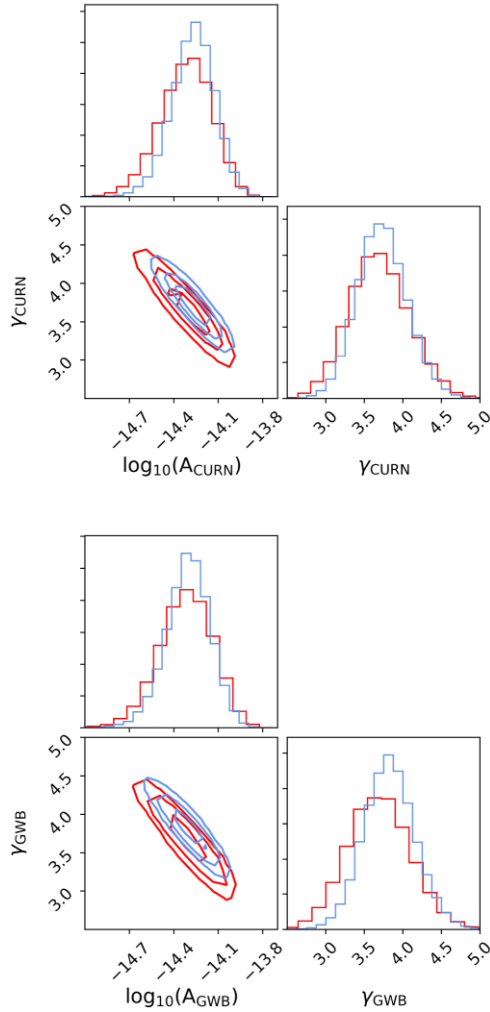
be CURN or GWB). Analysis with the custom noise model built in this work *confirms* these findings. It is quite remarkable that even though the custom noise models are very different from the standard ones, we can still confirm the presence of a CRS. This gives additional confidence in the presence of a CRS in the EPTA DR2 data.

The results presented in Table 7 should be considered more as a comparison of the nature of the CRS: PSRN+GWB ver-



**Figure 9.** Dropout score for the contribution of each pulsar to the CURN model with M1 (pink dots) or M2 (empty blue dots). The same analysis for PSR J1744–1134 but without intrinsic red-noise for this pulsar is also shown (empty red dot).

sus PSRN+CURN. It is worth mentioning that in this view the custom noise model M2 actually slightly prefers the GWB:  $\log_{10} \mathcal{B}_{\text{PSRN+GWB}}^{\text{PSRN+CURN}} = 0.3$ , whereas it is slightly negative ( $-0.2$ ) for M1 as estimated with the hypermodel method. The observed signal could also be caused by the spatially uncorrelated component of the GWB, which is expected to dominate over the cross-correlated terms.



**Figure 10.** 2D posterior distributions of the CURN (top) and Hellings–Down correlated GWB (bottom) power-law parameters with M1 (blue) and M2 (red) single-pulsar noise models.

**Table 6.** Median of amplitude and spectral index posterior (and 95 per cent credible interval) of both uncorrelated (CURN) and Hellings–Down (GWB) common processes computed with M1 and M2. The last two columns contain Jensen–Shannon divergences values that compare the posteriors of a CRS added to M1 and M2.

| Model   | $\log_{10}A$             | $\gamma$               | JS div. $A$           | JS div. $\gamma$      |
|---------|--------------------------|------------------------|-----------------------|-----------------------|
| M1+CURN | $-14.27^{+0.26}_{-0.32}$ | $3.73^{+0.68}_{-0.61}$ | 0                     | 0                     |
| M2+CURN | $-14.31^{+0.30}_{-0.38}$ | $3.68^{+0.81}_{-0.72}$ | $1.55 \times 10^{-2}$ | $9.40 \times 10^{-3}$ |
| M1+GWB  | $-14.30^{+0.27}_{-0.32}$ | $3.82^{+0.70}_{-0.65}$ | 0                     | 0                     |
| M2+GWB  | $-14.31^{+0.32}_{-0.39}$ | $3.69^{+0.84}_{-0.78}$ | $8.73 \times 10^{-3}$ | $2.60 \times 10^{-2}$ |

## 7 CONCLUSION

We used a general Bayesian inference approach to select the most favoured noise model for each pulsar of EPTA DR2. These models are summarized in Table 5 and show a significant improvement (in terms of Bayes factor) over the default base model used in Chen et al. (2021). In addition to conventional stochastic processes such as achromatic red noise and DM variations, we have considered scattering variations, system noise (both chromatic and achromatic),

band noise, and deterministic signals (annual DM variations, annual scattering variations, and non-stationary DM event). Our model selection was guided by previously published investigations or using auxiliary runs that helped to identify the list of models for further consideration.

The main result of this paper is that we confirm the presence of a common red noise in EPTA DR2 despite the use of much more complicated single-pulsar noise models. The data is not informative on the presence of Hellings–Downs spatial correlations, which is not surprising given that we have analysed only the six best EPTA pulsars. However, the use of custom noise models led to a marginal increase in the evidence for the presence of a GWB against an uncorrelated CRS.

The second main result of this paper is that it lays the practical scheme (protocol) for choosing a custom noise model which will be applied to a larger set of EPTA (and IPTA) pulsars on the path to confirming the nature of the observed common red noise process. Even though we only found small differences, both in posteriors and Bayes factors, the analysis of noise in each pulsar data is essential for the detection and interpretation of the common red signal.

There are also several interesting side results obtained during this investigation.

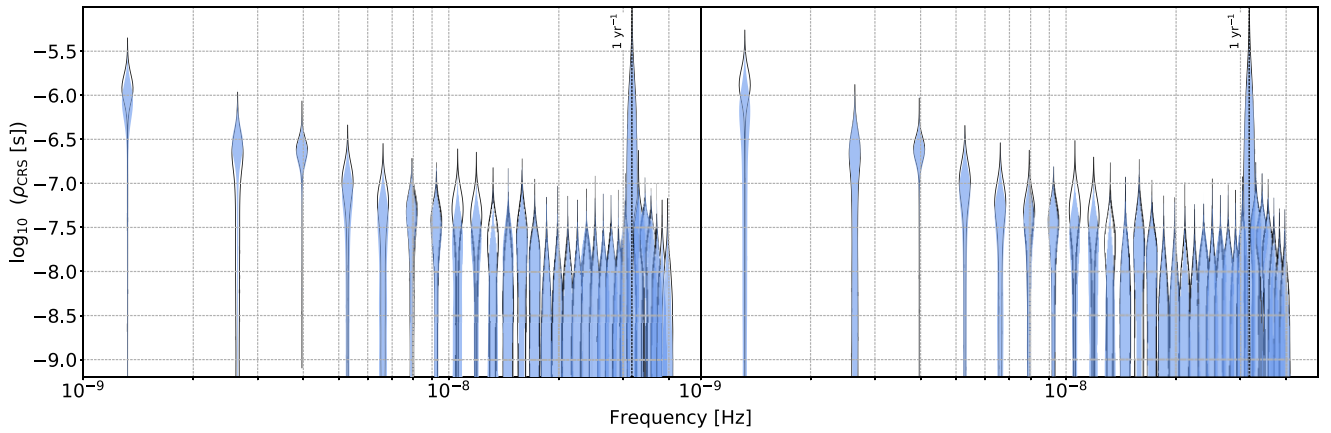
- (i) PSR J1909–3744 shows signs of a downturn in its spectrum at low frequencies. This feature could be confirmed or disproved with a longer observational span.
- (ii) PSR J1012+5307 indicates the presence of a high level of red noise at high frequencies of unclear origin.
- (iii) PSR J1600–3053 and J1744–1134 are not very informative about the presence of achromatic red noise, giving only a small preference to the models with RN.
- (iv) We did not observe strong evidence for the annual DM variations signal previously reported for PSR J0613–0200.
- (v) We found that the first exponential dip of PSR J1713+0747 has a chromatic index consistent with scattering variations, and confirmed the low chromatic index for the second event.

In this work we have emphasized several times the need to combine the data from multiple systems (to disentangle SN and RN processes) and multiband observations (to disentangle achromatic and chromatic red noise). This gives a strong reason for a joint analysis of combined IPTA data.

In the Appendix we show the red noise free spectrum with the default model *RN30\_DM100* and its time-domain realization for each pulsar. We also provide a Table B1 with the parameters of all noise components.

## ACKNOWLEDGEMENTS

The European Pulsar Timing Array (EPTA) is a collaboration between European and partner institutes, namely ASTRON (NL), INAF/Osservatorio di Cagliari (IT), Max-Planck-Institut für Radioastronomie (GER), Nançay/Paris Observatory (FRA), the University of Manchester (UK), the University of Birmingham (UK), the University of East Anglia (UK), the University of Bielefeld (GER), the University of Paris (FRA), the University of Milan-Bicocca (IT) and Peking University (CHN), with the aim to provide high precision pulsar timing to work towards the direct detection of low-frequency gravitational waves. An Advanced Grant of the European Research Council to implement the Large European Array for Pulsars (LEAP) also provides funding. The EPTA is part of the International Pulsar Timing Array (IPTA), we would like to thank our IPTA colleagues for their help with this paper.



**Figure 11.** Free-spectrum of the CURN (left) and the GWB (right) signals, either with M1 (empty black violin) or M2 (filled blue violin) single-pulsar noise models.

**Table 7.** Bayes factors in favour of the inclusion of CRS signals (CURN or GWB) in the M1 or M2 single-pulsar noise models. Estimations are performed either through evidence ratios (Dyn.) or using the product-space method (Hyp.).

| CRS Model | $\log_{10}$ BF     |                    |                    |                    |
|-----------|--------------------|--------------------|--------------------|--------------------|
|           | M1 <sub>Dyn.</sub> | M1 <sub>Hyp.</sub> | M2 <sub>Dyn.</sub> | M2 <sub>Hyp.</sub> |
| PSRN      | –                  | –                  | –                  | –                  |
| PSRN+CURN | 3.6                | 3.8                | 2.9                | 3.3                |
| PSRN+GWB  | 3.4                | 3.6                | 3.0                | 3.6                |

Part of this work is based on observations with the 100-m telescope of the Max-Planck-Institut für Radioastronomie (MPIfR) at Effelsberg in Germany. Pulsar research at the Jodrell Bank Centre for Astrophysics and the observations using the Lovell Telescope are supported by a Consolidated Grant (ST/T000414/1) from the UK’s Science and Technology Facilities Council. The Nançay radio Observatory is operated by the Paris Observatory, associated to the French Centre National de la Recherche Scientifique (CNRS), and partially supported by the Region Centre in France. We acknowledge financial support from ‘Programme National de Cosmologie et Galaxies’ (PNCG), and ‘Programme National Hautes Energies’ (PNHE) funded by CNRS/INSU-IN2P3-INP, CEA and CNES, France. We acknowledge financial support from Agence Nationale de la Recherche (ANR-18-CE31-0015), France. The Westerbork Synthesis Radio Telescope is operated by the Netherlands Institute for Radio Astronomy (ASTRON) with support from the Netherlands Foundation for Scientific Research (NWO). The Sardinia Radio Telescope (SRT) is funded by the Department of University and Research (MIUR), the Italian Space Agency (ASI), and the Autonomous Region of Sardinia (RAS) and is operated as National Facility by the National Institute for Astrophysics (INAF).

The work is supported by National SKA program of China 2020SKA0120100, Max-Planck Partner Group, NSFC 11690024, CAS Cultivation Project for FAST Scientific. This work is supported as part of the ‘LEGACY’ MPG-CAS collaboration on low-frequency gravitational wave astronomy. AC acknowledges support from the Paris Île-de-France Region. GS, AS, and AS acknowledge financial support provided under the European Union’s H2020 ERC Consolidator Grant ‘Binary Massive Black Hole Astrophysic’ (B

Massive, Grant Agreement: 818691). JA acknowledges support by the Stavros Niarchos Foundation (SNF) and the Hellenic Foundation for Research and Innovation (H.F.R.I.) under the 2nd Call of ‘Science and Society’ Action Always strive for excellence – ‘Theodoros Papazoglou’ (Project Number: 01431). GD, RK, and MK acknowledge support from European Research Council (ERC) Synergy Grant ‘BlackHoleCam’ Grant Agreement Number 610058 and ERC Advanced Grant ‘LEAP’ Grant Agreement Number 337062. JWM is a CITA Postdoctoral Fellow: This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), (funding reference #CITA 490888-16). AV acknowledges the support of the Royal Society and Wolfson Foundation. JPWV acknowledges support by the Deutsche Forschungsgemeinschaft (DFG) through the Heisenberg programme (Project No. 433075039). Finally, we would like to thank B. Goncharov for reading and commenting on the manuscript.

We used the following software packages to produce results and figures in this work: LIBSTEMPO (Vallisneri 2020), ENTERPRISE (Ellis et al. 2019), ENTERPRISE.EXTENSIONS ([https://github.com/nanogra/enterprise\\_extensions](https://github.com/nanogra/enterprise_extensions)), PTMCMCSAMPLER (Ellis & van Haasteren 2017), DYNESTY (Speagle 2020), MC<sup>3</sup> (<https://gitlab.in2p3.fr/stas/samplermcmc>), NUMPY (van der Walt, Colbert & Varoquaux 2011), SCIPY (Oliphant 2007), MATPLOTLIB (Hunter 2007), CORNER (Foreman-Mackey 2016), LA-FORGE (Hazboun 2020), and OUTLIER-UTILS (<https://github.com/c-bata/outlier-utils>).

## DATA AVAILABILITY

The timing data used in this article shall be shared on reasonable request to the corresponding authors.

## REFERENCES

- Alam M. F. et al., 2021, *ApJS*, 252, 4
- Anderson T. W., Darling D. A., 1952, *Ann. Math. Stat.*, 23, 193
- Arzoumanian Z. et al., 2020, *ApJ*, 905, L34
- Babak S. et al., 2015, *MNRAS*, 455, 1665
- Bassa C. G. et al., 2015, *MNRAS*, 456, 2196
- Brooks S. P., Gelman A., 1998, *J. Comput. Graph. Stat.*, 7, 434
- Caballero R. N. et al., 2016, *MNRAS*, 457, 4421
- Carlin B. P., Chib S., 1995, *J. R. Stat. Soc. Ser. B*, 57, 473
- Chen S., Sesana A., Del Pozzo W., 2017, *MNRAS*, 470, 1738



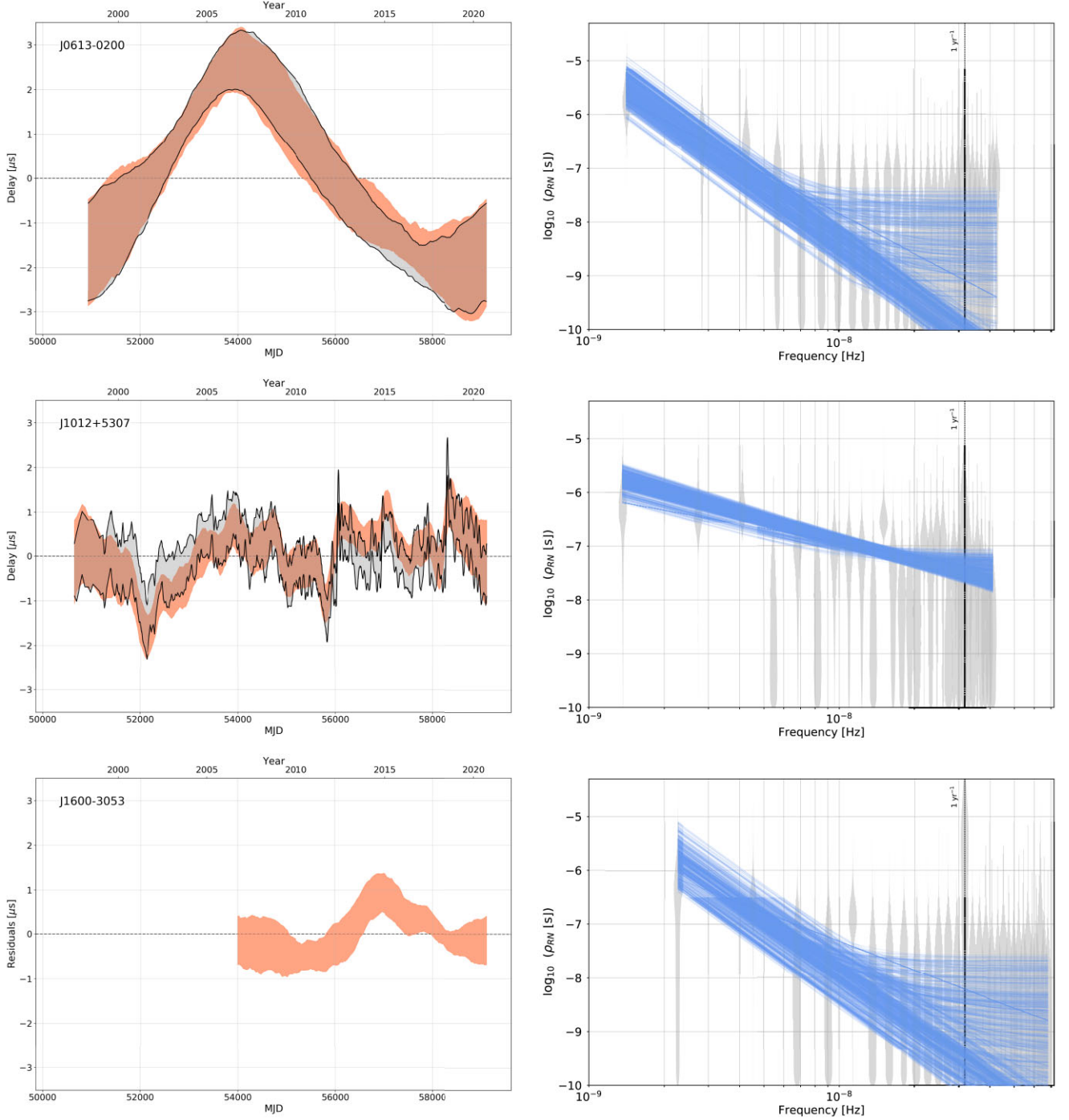
- Chen S. et al., 2021, *MNRAS*, 508, 4970
- Coles W. A. et al., 2015, *ApJ*, 808, 113
- Cordes J. M., Shannon R. M., 2010, preprint ([arXiv:1010.3785](https://arxiv.org/abs/1010.3785))
- Cordes J. M., Wolszczan A., 1986, *ApJ*, 307, L27
- Cordes J. M., Shannon R. M., Stinebring D. R., 2016, *ApJ*, 817, 16
- Desvignes G. et al., 2016, *MNRAS*, 458, 3341
- Detweiler S., 1979, *ApJ*, 234, 1100
- Ellis J., van Haasteren R., 2017, jellis18/ptmcmcsampler: Official Release. Available at <https://zenodo.org/record/1037579#.YXQLhpbZIU>
- Ellis J. A., Vallisneri M., Taylor S. R., Baker P. T., 2019, ENTERPRISE: Enhanced Numerical Toolbox Enabling a Robust Pulsar Inference Suite. Available at <https://zenodo.org/record/4059815#.YXQLuBpBzIU>
- Foreman-Mackey D., 2016, *J. Open Source Softw*, 1, 24
- Foster R. S., Backer D. C., 1990, *ApJ*, 361, 300
- Gelman A., Rubin D. B., 1992, *Stat. Sci.*, 7, 457
- Goncharov B., Zhu X.-J., Thrane E., 2020, *MNRAS*, 497, 3264
- Goncharov B. et al., 2021a, *ApJ*, 917, L19
- Goncharov B. et al., 2021b, *MNRAS*, 502, 478
- Grubbs F. E., 1950, *Ann. Math. Stat.*, 21, 27
- Hastings W. K., 1970, *Biometrika*, 57, 97
- Hazboun J. S., 2020, Hazboun6/la-forge. Available at <https://zenodo.org/record/4152550#.YaPWF5HJMjHe>
- Hazboun J. S., Simon J., Siemens X., Romano J. D., 2020, *ApJ*, 905, L6
- Hee S., Handley W. J., Hobson M. P., Lasenby A. N., 2015, *MNRAS*, 455, 2461
- Hellings R. W., Downs G. S., 1983, *ApJ*, 265, L39
- Hobbs G. B., Edwards R. T., Manchester R. N., 2006, *MNRAS*, 369, 655
- Hobbs G. et al., 2020, *MNRAS*, 491, 5951
- Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90
- Jaffe A. H., Backer D. C., 2003, *ApJ*, 583, 616
- Jeffreys H., 1961, *Theory of Probability*, 3rd edn. Clarendon Press, Oxford
- Jenet F. A. et al., 2006, *ApJ*, 653, 1571
- Kass R. E., Raftery A. E., 1995, *J. Am. Stat. Assoc.*, 90, 773
- Keith M. J. et al., 2013, *MNRAS*, 429, 2161
- Kerr M. et al., 2020, *PASA*, 37, e020
- Kullback S., 1959, *Information Theory and Statistics*. Wiley, New York
- Lam M. T. et al., 2016, *ApJ*, 819, 155
- Lam M. T. et al., 2018, *ApJ*, 861, 132
- Lee K. J., Jenet F. A., Price R. H., 2008, *ApJ*, 685, 1304
- Lentati L., Alexander P., Hobson M. P., Taylor S., Gair J., Balan S. T., van Haasteren R., 2013, *Phys. Rev. D*, 87, 104021
- Lentati L. et al., 2015, *MNRAS*, 453, 2576
- Lentati L. et al., 2016, *MNRAS*, 458, 2161
- Liu K., Keane E. F., Lee K. J., Kramer M., Cordes J. M., Purver M. B., 2012, *MNRAS*, 420, 361
- Lorimer D. R., Kramer M., 2004, *Handbook of Pulsar Astronomy*. Cambridge Univ. Press, Cambridge
- Lyne A., Hobbs G., Kramer M., Stairs I., Stappers B., 2010, *Science*, 329, 408
- Maggiore M., 2000, *Phys. Rep.*, 331, 283
- Main R. A. et al., 2020, *MNRAS*, 499, 1468
- Manning C. D., Schütze H., 1999, *Foundations of Statistical Natural Language Processing*. The MIT Press, Cambridge, <http://nlp.stanford.edu/fnlp/>
- Matsakis D. N., Taylor J. H., Eubanks T. M., 1997, *A&A*, 326, 924
- Melatos A., Link B., 2014, *MNRAS*, 437, 21
- Metropolis N., Rosenbluth A. W., Rosenbluth M. N., Teller A. H., Teller E., 1953, *J. Chem. Phys.*, 21, 1087
- Oliphant T. E., 2007, *Comput. Sci. Eng.*, 9, 10
- Perera B. B. P. et al., 2018, *MNRAS*, 478, 218
- Perera B. B. P. et al., 2019, *MNRAS*, 490, 4666
- Rajagopal M., Romani R. W., 1995, *ApJ*, 446, 543
- Reardon D. J. et al., 2016, *MNRAS*, 455, 1751
- Sazhin M. V., 1978, *Sov. Astron.*, 22, 36
- Sesana A., Haardt F., Madau P., Volonteri M., 2004, *ApJ*, 611, 623
- Sesana A., Vecchio A., Volonteri M., 2009, *MNRAS*, 394, 2255
- Shannon R. M., Cordes J. M., 2010, *ApJ*, 725, 1607
- Shannon R. M. et al., 2014, *MNRAS*, 443, 1463
- Sivia D. S., Skilling J., 2006, *Data Analysis - A Bayesian Tutorial*, 2nd edn. Oxford Univ. Press, Oxford
- Skilling J., 2004, in Fischer R., Preuss R., Toussaint U. V., eds, *AIP Conference Proc. Vol. 735, 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*. Am. Inst. Phys., New York, p. 395
- Skilling J., 2006, *Bayesian Anal.*, 1, 833
- Speagle J. S., 2020, *MNRAS*, 493, 3132
- Taylor J. H., 1992, *Phil. Trans. R. Soc. A*, 341, 117
- Taylor S. R., van Haasteren R., Sesana A., 2020, *Phys. Rev. D*, 102, 084039
- Tiburzi C. et al., 2015, *MNRAS*, 455, 4339
- Tsang D., Gourgoulatos K. N., 2013, *ApJ*, 773, L17
- Vallisneri M., 2013, *vallis/libstempo*. Available at <https://github.com/vallis/libstempo/>
- van Haasteren R., Levin Y., 2012, *MNRAS*, 428, 1147
- van Haasteren R., Vallisneri M., 2014, *Phys. Rev. D*, 90, 104012
- van Haasteren R., Vallisneri M., 2015, *MNRAS*, 446, 1170
- van Haasteren R., Levin Y., McDonald P., Lu T., 2009, *MNRAS*, 395, 1005
- van Straten W., 2013, *ApJS*, 204, 13
- van der Walt S., Colbert S. C., Varoquaux G., 2011, *Comput. Sci. Eng.*, 13, 22
- Verbiest J. P. W. et al., 2009, *MNRAS*, 400, 951
- Verbiest J. P. W. et al., 2016, *MNRAS*, 458, 1267
- Wyithe J. S. B., Loeb A., 2003, *ApJ*, 590, 691
- You X. P. et al., 2007, *MNRAS*, 378, 493
- Zhu W. W. et al., 2015, *ApJ*, 809, 41

## APPENDIX A: ACHROMATIC RED-NOISE PROPERTIES

Here we provide additional plots which demonstrate our main findings about the RN in each pulsar.

We reconstruct time-domain Gaussian process realizations of this signal modelled with a power-law PSD. The left-hand panels of Fig. A1 display the 68 per cent confidence interval of 100 random realizations drawn from the posterior distributions of the RN amplitude and spectral index included in two different single-pulsar noise models: the default base model, *RN30\_DMv100* (red), and the ‘custom’ model shown in Table 5 (grey). The peculiar high frequency red noise is clearly seen in PSR J1012+5307 (see Section 5.2 for a detailed discussion). The RN in PSR J1744–1134 is considerably reduced in the ‘custom’ model (see discussion at the end of Section 5.2). Note that we did not include RN in the custom model for PSR J1600–3053. The RN in the default and custom models are quite similar for PSRs J0613–0200, J1713+0747, and J1909–3053.

The right-hand panels of Fig. A1 display the spectrum of the achromatic red-noise (using *RN30\_DMv100*) for each pulsar computed with (i) a free-spectrum PSD (grey violins); and (ii) a broken power-law PSD (blue), where we give 1000 realizations randomly drawn from the posterior distributions.



**Figure A1.** (left-hand panels) 68 per cent confidence interval of 100 time-domain random realizations of the achromatic red-noise included in the default base (*RN30\_DM100*) model (red) and the ‘custom’ selected models (light grey) (cf. Table 5) for each of the six pulsars. (right-hand panels) Achromatic red-noise spectrum for the corresponding pulsar (noted in left-hand panel) included in the default base model and described with (i) a free-spectrum PSD (grey violins); or (ii) a broken power law (blue), here showing 1000 random realizations drawn from the posterior distributions.

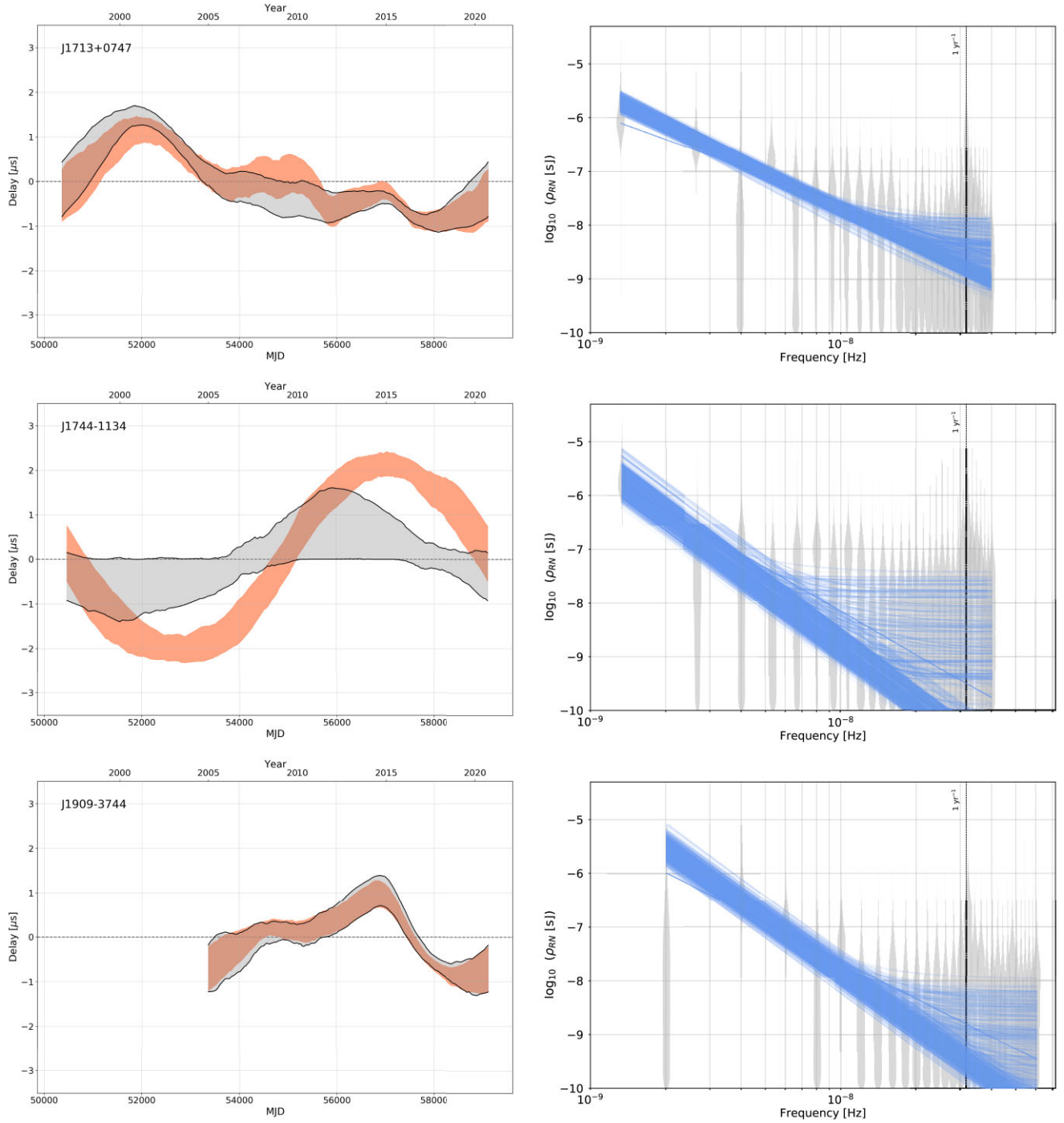


Figure A1 – continued

## APPENDIX B: SINGLE-PULSAR NOISE MODEL PARAMETERS

In Table B1 we give median values for each noise component with 68 per cent confidence interval in the custom model of each pulsar.

**Table B1.** Medians and 16–84 per cent credible intervals of the 1D marginalized posterior distributions of each single-pulsar noise model parameters for (i) the default base models used in Chen et al. (2021); and (ii) the final custom models. Note that  $DM\nu$  amplitudes are given with ENTERPRISE normalization set for the radio frequency at 1.4 GHz (see equation 8).

| Model   | Signal              | Parameter             | J0613–0200               | J1012+5307               | J1600–3053               | J1713+0747                 | J1744–1134               | J1909–3744               |
|---------|---------------------|-----------------------|--------------------------|--------------------------|--------------------------|----------------------------|--------------------------|--------------------------|
| Default | RN                  | $\log_{10}A$          | $-14.72^{+0.56}_{-0.60}$ | $-13.12^{+0.08}_{-0.08}$ | $-14.05^{+0.33}_{-0.53}$ | $-14.13^{+0.18}_{-0.19}$   | $-15.16^{+0.69}_{-0.72}$ | $-14.65^{+0.32}_{-0.37}$ |
|         |                     | $\gamma$              | $4.76^{+1.09}_{-1.04}$   | $1.66^{+0.32}_{-0.30}$   | $3.50^{+1.22}_{-0.89}$   | $3.29^{+0.54}_{-0.47}$     | $5.19^{+1.10}_{-1.09}$   | $4.65^{+0.96}_{-0.83}$   |
|         | $DM\nu$             | $\log_{10}A$          | $-13.75^{+0.21}_{-0.26}$ | $-13.36^{+0.05}_{-0.05}$ | $-13.09^{+0.04}_{-0.04}$ | $-13.47^{+0.04}_{-0.04}$   | $-13.33^{+0.05}_{-0.05}$ | $-13.56^{+0.04}_{-0.04}$ |
|         |                     | $\gamma$              | $2.89^{+0.66}_{-0.59}$   | $1.20^{+0.18}_{-0.16}$   | $2.08^{+0.12}_{-0.11}$   | $1.49^{+0.20}_{-0.19}$     | $1.23^{+0.21}_{-0.21}$   | $1.53^{+0.15}_{-0.14}$   |
| Custom  | RN                  | $\log_{10}A$          | $-14.82^{+0.64}_{-0.67}$ | $-13.02^{+0.03}_{-0.03}$ | –                        | $-14.48^{+0.25}_{-0.25}$   | $-15.25^{+1.31}_{-1.39}$ | $-14.46^{+0.36}_{-0.41}$ |
|         |                     | $\gamma$              | $4.81^{+1.17}_{-1.15}$   | $1.19^{+0.13}_{-0.13}$   | –                        | $3.95^{+0.65}_{-0.56}$     | $3.67^{+2.17}_{-2.22}$   | $4.24^{+1.03}_{-0.88}$   |
|         | $DM\nu$             | $\log_{10}A$          | $-13.58^{+0.16}_{-0.21}$ | $-13.66^{+0.12}_{-0.14}$ | $-14.16^{+0.32}_{-0.41}$ | $-13.78^{+0.05}_{-0.06}$   | $-13.45^{+0.07}_{-0.07}$ | $-13.92^{+0.27}_{-1.21}$ |
|         |                     | $\gamma$              | $2.47^{+0.55}_{-0.47}$   | $2.09^{+0.42}_{-0.39}$   | $4.69^{+0.92}_{-0.76}$   | $1.16^{+0.20}_{-0.21}$     | $0.46^{+0.36}_{-0.30}$   | $2.64^{+2.67}_{-0.93}$   |
|         | Sv                  | $\log_{10}A$          | –                        | –                        | $-13.26^{+0.04}_{-0.04}$ | –                          | –                        | $-13.84^{+0.10}_{-0.19}$ |
|         |                     | $\gamma$              | –                        | –                        | $1.48^{+0.14}_{-0.13}$   | –                          | –                        | $0.78^{+0.31}_{-0.34}$   |
|         | Exp. dip 1          | $\log_{10}A$ [s]      | –                        | –                        | –                        | $-5.54^{+0.05}_{-0.04}$    | –                        | –                        |
|         |                     | $\log_{10}\tau$ [day] | –                        | –                        | –                        | $1.54^{+0.07}_{-0.07}$     | –                        | –                        |
|         |                     | $t_0$ [MJD]           | –                        | –                        | –                        | $54752.49^{+3.24}_{-3.17}$ | –                        | –                        |
|         | Exp. dip 2          | $\log_{10}A$ [s]      | –                        | –                        | –                        | $-5.89^{+0.05}_{-0.05}$    | –                        | –                        |
|         |                     | $\log_{10}\tau$ [day] | –                        | –                        | –                        | $1.51^{+0.09}_{-0.09}$     | –                        | –                        |
|         |                     | $t_0$ [MJD]           | –                        | –                        | –                        | $57510.65^{+2.17}_{-2.20}$ | –                        | –                        |
|         | SN_BON_2.0          | $\log_{10}A$          | –                        | –                        | –                        | $-14.78^{+1.03}_{-0.97}$   | –                        | –                        |
|         |                     | $\gamma$              | –                        | –                        | –                        | $4.36^{+1.59}_{-1.80}$     | –                        | –                        |
|         | SN_JBO_1.5          | $\log_{10}A$          | –                        | –                        | –                        | $-13.10^{+0.21}_{-0.22}$   | –                        | –                        |
|         |                     | $\gamma$              | –                        | –                        | –                        | $1.47^{+0.68}_{-0.65}$     | –                        | –                        |
|         | SN_LEAP_1.4         | $\log_{10}A$          | –                        | –                        | $-14.97^{+0.94}_{-0.66}$ | $-13.42^{+0.20}_{-0.21}$   | –                        | –                        |
|         |                     | $\gamma$              | –                        | –                        | $5.16^{+1.19}_{-1.75}$   | $1.70^{+0.60}_{-0.59}$     | –                        | –                        |
|         | $DM\nu$ -SN_NUP_1.4 | $\log_{10}A$          | $-13.52^{+0.29}_{-0.33}$ | $-14.10^{+0.72}_{-0.91}$ | –                        | $-14.05^{+0.63}_{-0.64}$   | $-14.65^{+0.99}_{-1.03}$ | –                        |
|         |                     | $\gamma$              | $2.88^{+0.95}_{-0.85}$   | $3.17^{+1.60}_{-1.31}$   | –                        | $2.75^{+1.20}_{-1.14}$     | $4.23^{+1.68}_{-1.75}$   | –                        |
|         | SN_NUP_2.5          | $\log_{10}A$          | –                        | $-13.22^{+0.73}_{-1.03}$ | –                        | –                          | –                        | –                        |
|         |                     | $\gamma$              | –                        | $2.50^{+1.84}_{-1.44}$   | –                        | –                          | –                        | –                        |
|         | BN_Band.2           | $\log_{10}A$          | –                        | –                        | –                        | –                          | $-13.84^{+0.35}_{-0.42}$ | –                        |
|         |                     | $\gamma$              | –                        | –                        | –                        | –                          | $3.02^{+0.86}_{-0.77}$   | –                        |
|         | BN_Band.3           | $\log_{10}A$          | –                        | –                        | –                        | $-14.32^{+0.37}_{-0.48}$   | –                        | –                        |
|         |                     | $\gamma$              | –                        | –                        | –                        | $2.68^{+1.02}_{-0.87}$     | –                        | –                        |

<sup>1</sup>CNRS, Astroparticule et Cosmologie, Université de Paris, F-75013 Paris, France

<sup>2</sup>Laboratoire de Physique et Chimie de l'Environnement et de l'Espace LPC2E UMR7328, CNRS, Université d'Orléans, F-45071 Orléans, France

<sup>3</sup>Station de Radioastronomie de Nançay, Observatoire de Paris, PSL University, CNRS, Université d'Orléans, F-18330 Nançay, France

<sup>4</sup>Moscow Institute of Physics and Technology, Dolgoprudny, Moscow region, Moscow 141700, Russia

<sup>5</sup>IRFU, CEA, Université Paris-Saclay, F-91191, Gif-sur-Yvette, France

<sup>6</sup>Dipartimento di Fisica ‘G. Occhialini’, Università degli Studi di Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy

<sup>7</sup>Kavli Institute for Astronomy and Astrophysics, Peking University, 100871 Beijing, P. R. China

<sup>8</sup>Laboratoire Univers et Théories LUTH, Observatoire de Paris, Université PSL, CNRS, Université de Paris, F-92190 Meudon, France

<sup>9</sup>Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany

<sup>10</sup>LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de Paris, 5 place Jules Janssen, F-92195 Meudon, France

<sup>11</sup>INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy

<sup>12</sup>ASTRON, Netherlands Institute for Radio Astronomy, Oude Hoogeveensedijk 4, NL-7991 PD Dwingeloo, the Netherlands

<sup>13</sup>Department of Astrophysics/IMAPP, Radboud University Nijmegen, P.O. Box 9010, NL-6500 GL Nijmegen, the Netherlands

<sup>14</sup>Institute of Astrophysics, FORTH, N. Plastira 100, 70013, Heraklion, Greece

<sup>15</sup>Argelander Institut für Astronomie, Auf dem Hügel 71, D-53117 Bonn, Germany

<sup>16</sup>Fakultät für Physik, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Germany

<sup>17</sup>INAF - Osservatorio Astronomico di Cagliari, via della Scienza 5, I-09047 Selargius (CA), Italy



<sup>18</sup>*School of Physics, Faculty of Science, University of East Anglia, NR4 7TJ Norwich, UK*

<sup>19</sup>*Max-Planck-Institut für Gravitationsphysik (Albert Einstein Institut), Am Mühlenberg 1, D-14476 Golm, Germany*

<sup>20</sup>*Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, University of Manchester, M13 9PL Manchester, UK*

<sup>21</sup>*National Astronomical Observatories, Chinese Academy of Sciences, 100101 Beijing, P. R. China*

<sup>22</sup>*Laboratory of Gravitational Waves and Cosmology, Advanced Institute of Natural Sciences, Normal University at Zhuhai, 519087 Beijing, P. R. China*

<sup>23</sup>*Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, ON M5S 3H8 Toronto, Canada*

<sup>24</sup>*Arecibo Observatory, HC3 Box 53995, PR 00612 Arecibo, USA*

<sup>25</sup>*Max-Planck-Institut für Gravitationsphysik, Albert Einstein Institut, Am Mühlenberg 1, D-14476 Golm, Germany*

<sup>26</sup>*Institute for Gravitational Wave Astronomy and School of Physics and Astronomy, University of Birmingham, Edgbaston, B15 2TT Birmingham, UK*

<sup>27</sup>*Department of Astronomy, Peking University, Beijing 100871, P. R. China*

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