

# Practical approaches to analyzing PTA data: Cosmic strings with six pulsars

Hippolyte Quelquejay Leclerc<sup>1,\*</sup>, Pierre Auclair<sup>2,†</sup>, Stanislav Babak<sup>1,‡</sup>, Aurélien Chalumeau<sup>3</sup>, Danièle A. Steer<sup>1,§</sup>, J. Antoniadis<sup>4,5</sup>, A.-S. Bak Nielsen<sup>5,6</sup>, C. G. Bassa<sup>7</sup>, A. Berthereau<sup>8,9</sup>, M. Bonetti<sup>3,10,11</sup>, E. Bortolas<sup>3,10,11</sup>, P. R. Brook<sup>12</sup>, M. Burgay<sup>13</sup>, R. N. Caballero<sup>14</sup>, D. J. Champion<sup>5</sup>, S. Chanlaridis<sup>4</sup>, S. Chen<sup>15</sup>, I. Cognard<sup>8,9</sup>, G. Desvignes<sup>5</sup>, M. Falxa<sup>8,1</sup>, R. D. Ferdman<sup>16</sup>, A. Franchini<sup>3,10</sup>, J. R. Gair<sup>17</sup>, B. Goncharov<sup>18,19</sup>, E. Graikou<sup>5</sup>, J.-M. Grießmeier<sup>8,9</sup>, L. Guillemot<sup>8,9</sup>, Y. J. Guo<sup>5</sup>, H. Hu<sup>5</sup>, F. Iraci<sup>13,20</sup>, D. Izquierdo-Villalba<sup>3,10</sup>, J. Jang<sup>5</sup>, J. Jawor<sup>5</sup>, G. H. Janssen<sup>7,21</sup>, A. Jessner<sup>5</sup>, R. Karuppusamy<sup>5</sup>, E. F. Keane<sup>22</sup>, M. J. Keith<sup>23</sup>, M. Kramer<sup>5,23</sup>, M. A. Krishnakumar<sup>5,6</sup>, K. Lackeos<sup>5</sup>, K. J. Lee<sup>4,5,9</sup>, K. Liu<sup>5</sup>, Y. Liu<sup>6,24</sup>, A. G. Lyne<sup>23</sup>, J. W. McKee<sup>25,26</sup>, R. A. Main<sup>5</sup>, M. B. Mickaliger<sup>23</sup>, I. C. Nițu<sup>23</sup>, A. Parthasarathy<sup>5</sup>, B. B. P. Perera<sup>27</sup>, D. Perrodin<sup>13</sup>, A. Petiteau<sup>28,1</sup>, N. K. Porayko<sup>5,3</sup>, A. Possenti<sup>13</sup>, A. Samajdar<sup>29</sup>, S. A. Sanidas<sup>23</sup>, A. Sesana<sup>3,10,11</sup>, G. Shaifullah<sup>3,10,13</sup>, L. Spери<sup>17</sup>, R. Spiewak<sup>23</sup>, B. W. Stappers<sup>23</sup>, S. C. Susarla<sup>30</sup>, G. Theureau<sup>8,9,31</sup>, C. Tiburzi<sup>13</sup>, E. van der Wateren<sup>21,7</sup>, A. Vecchio<sup>12</sup>, V. Venkatraman Krishnan<sup>5</sup>, J. Wang<sup>6,32,33</sup>, L. Wang<sup>23</sup> and Z. Wu<sup>24</sup>

(European Pulsar Timing Array)

<sup>1</sup>Université Paris Cité CNRS, Astroparticule et Cosmologie, 75013 Paris, France

<sup>2</sup>Cosmology, Universe and Relativity at Louvain (CURL), Institute of Mathematics and Physics, University of Louvain, 2 Chemin du Cyclotron, 1348 Louvain-la-Neuve, Belgium

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

<sup>4</sup>Institute of Astrophysics, FORTH, N. Plastira 100, 70013, Heraklion, Greece

<sup>5</sup>Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany

<sup>6</sup>Fakultät für Physik, Universität Bielefeld, Postfach 100131, 33501 Bielefeld, Germany

<sup>7</sup>ASTRON, Netherlands Institute for Radio Astronomy, Oude Hoogeveensedijk 4, 7991 PD, Dwingeloo, The Netherlands

<sup>8</sup>Laboratoire de Physique et Chimie de l’Environnement et de l’Espace, Université d’Orléans/CNRS, 45071 Orléans Cedex 02, France

<sup>9</sup>Observatoire Radioastronomique de Nançay, Observatoire de Paris, Université PSL, Université d’Orléans, CNRS, 18330 Nançay, France

<sup>10</sup>INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy

<sup>11</sup>INAF—Osservatorio Astronomico di Brera, via Brera 20, I-20121 Milano, Italy

<sup>12</sup>Institute for Gravitational Wave Astronomy and School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom

<sup>13</sup>INAF—Osservatorio Astronomico di Cagliari, via della Scienza 5, 09047 Selargius (CA), Italy

<sup>14</sup>Hellenic Open University, School of Science and Technology, 26335 Patras, Greece

<sup>15</sup>Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, People’s Republic of China

<sup>16</sup>School of Physics, Faculty of Science, University of East Anglia, Norwich NR4 7TJ, United Kingdom

<sup>17</sup>Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Am Mühlenberg 1, 14476 Potsdam, Germany

<sup>18</sup>Gran Sasso Science Institute (GSSI), I-67100 L’Aquila, Italy

<sup>19</sup>INFN, Laboratori Nazionali del Gran Sasso, I-67100 Assergi, Italy

<sup>20</sup>Università di Cagliari, Dipartimento di Fisica, S.P. Monserrato-Sestu Km 0,700–09042 Monserrato (CA), Italy

<sup>21</sup>Department of Astrophysics/IMAPP, Radboud University Nijmegen, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands

<sup>22</sup>School of Physics, Trinity College Dublin, College Green, Dublin 2, D02 PN40, Ireland

<sup>23</sup>Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, University of Manchester, Manchester M13 9PL, United Kingdom

<sup>24</sup>National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, People’s Republic of China

<sup>25</sup>E.A. Milne Centre for Astrophysics, University of Hull,

Cottingham Road, Kingston-upon-Hull, HU6 7RX, United Kingdom

<sup>26</sup>Centre of Excellence for Data Science, Artificial Intelligence and Modelling (DAIM), University of Hull, Cottingham Road, Kingston-upon-Hull, HU6 7RX, United Kingdom

<sup>27</sup>Arecibo Observatory, HC3 Box 53995, Arecibo, Puerto Rico 00612

<sup>28</sup>IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France

<sup>29</sup>*Institut für Physik und Astronomiew, Universität Potsdam,  
Haus 28, Karl-Liebknecht-Str. 24/25, 14476, Potsdam, Germany*

<sup>30</sup>*Ollscoil na Gaillimhe—University of Galway, University Road, Galway, H91 TK33, Ireland*

<sup>31</sup>*Laboratoire Univers et Théories LUTH, Observatoire de Paris, Université PSL,  
CNRS, Université de Paris, 92190 Meudon, France*

<sup>32</sup>*Ruhr University Bochum, Faculty of Physics and Astronomy,  
Astronomical Institute (AIRUB), 44780 Bochum, Germany*

<sup>33</sup>*Advanced Institute of Natural Sciences, Beijing Normal University, Zhuhai 519087, China*



(Received 1 July 2023; accepted 14 November 2023; published 15 December 2023)

We search for a stochastic gravitational wave background (SGWB) generated by a network of cosmic strings using six millisecond pulsars from Data Release 2 (DR2) of the European Pulsar Timing Array (EPTA). We perform a Bayesian analysis considering two models for the network of cosmic string loops, and compare it to a simple power-law model which is expected from the population of supermassive black hole binaries. Our main strong assumption is that the previously reported common red noise process is a SGWB. We find that the one-parameter cosmic string model is slightly favored over a power-law model thanks to its simplicity. If we assume a two-component stochastic signal in the data (supermassive black hole binary population and the signal from cosmic strings), we get a 95% upper limit on the string tension of  $\log_{10}(G\mu) < -9.9$  ( $-10.5$ ) for the two cosmic string models we consider. In extended two-parameter string models, we were unable to constrain the number of kinks. We test two approximate and fast Bayesian data analysis methods against the most rigorous analysis and find consistent results. These two fast and efficient methods are applicable to all SGWBs, independent of their source, and will be crucial for analysis of extended datasets.

DOI: [10.1103/PhysRevD.108.123527](https://doi.org/10.1103/PhysRevD.108.123527)

## I. INTRODUCTION

All regional pulsar timing array (PTA) collaborations have recently, and independently, reported the presence of a common red noise process in the observed data [1–3]. Furthermore, the combined dataset shows even stronger evidence for its presence [4]. The exact nature of this signal is as yet very uncertain. While its spectral properties are consistent with an expected stochastic gravitational wave background (SGWB), the data is not sensitive enough to make any informative statement for or against its gravitational wave (GW) nature.<sup>1</sup>

In this paper, we make a *strong* assumption that it is, in fact, a SGWB and work toward its possible interpretation. The most favorable model for an anticipated SGWB is a superposition of monochromatic GW signals from a population of supermassive black hole binaries (SMBHBs) in the local Universe [1,10]. This signal, assuming SMBHBs in circular orbits, is a power-law with theoretical spectral index  $\gamma = 13/3$  in the power spectral density of the residuals [11].

However, realistic astrophysical simulations suggest that the spectral index could vary based on the realization of the observed Universe [12]. For this reason, the spectral index is usually inferred from the data, giving a two-parameter model:  $A, \gamma$ . The amplitude  $A$  is referenced at the frequency 1/year.

We propose an alternative to SMBHBs, namely a SGWB from an early universe source, and in particular a network of cosmic string loops. (See [13–17] for previous work on cosmic strings in the nHz band.) Cosmic strings are topological defects that could have emerged from symmetry-breaking phase transitions in the early Universe [18–21]. These quasi one-dimensional objects are characterized by their dimensionless tension  $G\mu$ , where  $G$  is Newton’s constant. Numerical simulations show that on large scales, cosmic string networks reach an attractor “scaling” regime in which all the characteristic length scales in the network grow as  $t$  [22–25]. Cosmic string loops are formed at all times by the scaling infinite string network, oscillate with period  $\ell/2$ , and decay into GWs.<sup>2</sup> The superposition of the GW emitted generates a SGWB that depends on the cosmic string loop distribution. While the loop distribution is well understood on large scales close to the Hubble

\*quelquejay@apc.in2p3.fr

†pierre.auclair@uclouvain.be

‡stas@apc.in2p3.fr

§steer@apc.in2p3.fr

<sup>1</sup>After submission of this paper on the ArXiv and during the peer review process, the results from the regional PTA collaborations became public [5–9]. These papers have demonstrated evidence of GW-induced correlations in their respective datasets with a considerable statistical significance.

<sup>2</sup>Field theory simulations show that cosmic string loops also decay into particles [26,27]. Although the balance between GW and particle emission is still under debate, preliminary quantitative studies tend to show that the SGWB from cosmic string is unaffected at the frequency of PTAs [28,29].

TABLE I. Noise model used for each of six pulsars. For each pulsar, we also consider white noise whose parameters are set to the maximum likelihood values obtained through single pulsar analysis. We indicate the number of Fourier bins [ $N_f$  in Eq. (9)] employed for each noise type and pulsar. In cases where a particular noise type is not included in the pulsar’s noise model, we denote it with a  $\dots$  sign.

Pulsar	Time span (years)	Number of ToAs	Red noise	DM variations	Scattering variations	Deterministic signals
J0613 – 0200	22.89	2909	10	144	$\dots$	$\dots$
J1012 + 5307	23.68	5325	149	45	$\dots$	$\dots$
J1600 – 3053	14.32	2982	$\dots$	26	137	$\dots$
J1713 + 0747	24.46	5003	11	148	$\dots$	Two exponential dips <sup>a</sup>
J1744 – 1134	24.01	1946	9	151	$\dots$	$\dots$
J1909 – 3744	15.74	2503	20	151	$\dots$	$\dots$

<sup>a</sup>Interested readers are directed to Section 5.2.2 of [35] for further details regarding these events.

radius, there is an active debate regarding the distribution of smaller loops. To account for this theoretical uncertainty, in the following we consider two loop distribution models that have been used by the LIGO-Virgo-Kagra [30,31] and LISA [32,33] collaborations.

In this paper we use the six pulsars (see Table I) from the early Data Release 2 (DR2) of European PTA (EPTA) collaboration [1]<sup>3</sup>. Six pulsars, at the current sensitivity and observational time span, are not sufficient to detect the stochastic GW signal, however the presence of the common red process is evident. More pulsars are required to populate the expected Hellings-Downs (HD) [34] correlation curve and to get a high statistical significance for a SGWB. Nevertheless, in our study, we still utilize Hellings-Downs correlations to analyze the common red noise among pulsars. This approach allows us to account for the correlations among pulsars when characterizing the common red noise (CRN), resulting in a distinct interaction with the individual red noises. Given that we consider only 6 pulsars, the diagonal noise component (autocorrelation part of the noise matrix) will dominate and we expect only a small difference from the inclusion of Hellings-Downs correlation. Here we describe in detail the methodology of inferring parameters of the string network. This same methodology will later be applied on the extended dataset [5].

Assuming that the observed common red noise is a SGWB, we use Bayesian methods to infer the characteristics of the string network using a one-parameter model ( $G\mu$ ) and a two-parameter model ( $G\mu, N_k$ ) where  $N_k$  is the average number of kinks on a loop per oscillation period, see Sec. II. In addition to the SGWB signal, we also model the three components of the noise: white noise, pulsar red noise and chromatic noises (dispersion measurement variations and scattering variations) (see [35] for details).

We carry the Bayesian analysis through three different approaches with varying accuracy of description and computational cost. The first approach, dubbed in this

paper as the “full method” [36,37], is the standard approach in which the parameters of the SGWB and of the noise are explored simultaneously. This method is the most accurate but also the most computationally expensive due to the high-dimensionality of the analysis and because we consider Hellings-Downs correlations for the SGWB. The second approach, dubbed as the “resampling method” [38], neglects at first the spatial correlation between pulsars considering a common uncorrelated signal. The resulting posterior distributions are then resampled by taking into account GW-induced correlation. In this approach, the likelihood is factorized into a product of likelihoods for each pulsar [39], and the computational cost is significantly reduced with respect to the full method as the number of pulsars increases. Finally, the third method, dubbed as “free spectrum” [40,41], amounts to obtaining the correlated power of the common noise for each frequency bin independently, marginalizing over the single pulsar noise parameters, *before* inferring the SGWB parameters. This method drastically reduces the dimensionality of the analysis and allows for a very fast parameter estimation of the SGWB models.

The paper is organized as follows. We start with a description of the SGWB produced by cosmic string loops in Sec. II. In Sec. III we explain the three data analysis approaches used. In Sec. IV, we present the constraints on single parameter models ( $G\mu$ ) and two parameter models ( $G\mu, N_k$ ). We perform a Bayesian model comparison between the different cosmic string models and a simple power-law SGWB. Finally, we conclude in Sec. V.

## II. STOCHASTIC GRAVITATIONAL WAVE BACKGROUND FROM COSMIC STRINGS

The SGWB generated by a cosmic string network has been studied in depth, see, e.g., [32,42]. It includes a contribution from the uncorrelated superposition GW bursts emitted from cusps, kinks, and kink-kink collisions on oscillating loops [43]. Cusps, which are points at which the loop reaches ultra relativistic velocities, emit a short beam of GWs.

<sup>3</sup>This dataset is a subset of the full DR2 dataset recently used in [5] which includes a total of 25 pulsars.

Kinks, which are discontinuities in the tangent vector of the loop, propagate along the string at the speed of light and emit a one-dimensional “fanlike” GW burst with the same beaming angle as for cusps [43]. Finally, when left- and right-moving kinks collide (kink-kink collisions), an isotropic burst of GW emission is emitted [44].

A loop of length  $\ell$  oscillates periodically with corresponding fundamental frequency  $f_0 = 2/\ell$ . For the  $n$ th oscillatory mode of the loop, the power (in units of  $G\mu^2$ ) emitted in GWs by each of these burst events (labeled by  $b$ ) is given by<sup>4</sup> [43]

$$P_n^{(b)} = \frac{\Gamma^{(b)}}{\zeta(q_b)} n^{-q_b}. \quad (1)$$

Here  $\zeta(q_b)$  is the Riemann zeta function,  $q_b = 4/3, 5/3$ , and 2 for cusps, kinks and kink-kink collisions respectively and [43,44]

$$\Gamma^c = \frac{3(\pi g_1^c)^2}{2^{1/3} g_2^{2/3}}, \quad \Gamma^k = \frac{3(\pi g_1^k)^2}{2^{2/3} g_2^{1/3}}, \quad \Gamma^{kk} = 2(\pi g_1^{kk})^2, \quad (2)$$

where  $g_2 = \sqrt{3}/4$ ,  $g_1^c \approx 0.85$ ,  $g_1^k \approx 0.29$ ,  $g_1^{kk} \approx 0.1$ .

Given the exponents  $q_b$  in Eq. (1), one might expect that, compared to the cusp contribution, the kink and kink-kink collision contributions could be neglected. However, as already noted in [44], if many kinks are present they can dominate the power emitted in GWs. Indeed, since two kinks (right and left moving) are formed when two pieces of strings intercommute, the number of kink-kink collisions on a loop with  $N_k$  kinks per oscillation period should scale as  $N_k^2/4$ , which can then dominate for high  $N_k$ . Furthermore, kinks can proliferate on loops with junctions [45,46]. To summarize, for a loop containing  $N_c$  cusps and  $N_k$  kinks per oscillation period, the total power emitted, in units of  $G\mu^2$ , is

$$\Gamma = N_c \Gamma^c + N_k \Gamma^k + \frac{N_k^2}{4} \Gamma^{kk}. \quad (3)$$

The fractional energy density of the SGWB per logarithmic interval of frequency is

$$\Omega_{\text{gw}}(t_0, f) \equiv \frac{8\pi G}{3H_0^2} f \frac{d\rho_{\text{gw}}}{df}, \quad (4)$$

where  $H_0$  is the Hubble constant, and  $d\rho_{\text{gw}}/df$  is the energy density in gravitational waves per unit frequency  $f$ , observed today (at  $t = t_0$ ). Following [31,47], for a network of cosmic string loops with distribution  $\mathbf{n}(\ell, t)$ , this is given by

<sup>4</sup>Formally this expression is only valid for large  $n$ , but as elsewhere we assume that extrapolating to  $n \geq 1$  is a good approximation to describe low-harmonic modes.

$$\frac{d\rho_{\text{gw}}}{df} = \frac{2G\mu^2}{f} \sum_b \frac{N_b \Gamma^{(b)}}{\zeta(q_b)} \times \sum_{n=1}^{+\infty} \int \frac{n^{1-q_b} dz}{(1+z)^5 H(z)} \mathbf{n} \left[ \frac{2n}{(1+z)f}, t(z) \right], \quad (5)$$

where we have summed over the cusp, kink and kink-kink contributions, and  $H(z)$  is the Hubble parameter. In the following we consider standard  $\Lambda$ CDM cosmology with the Planck-2018 fiducial parameters [48].

We consider the most updated loop distribution models  $\mathbf{n}(\ell, t)$  available, calibrated with numerical simulations, that have also been used in LVK collaboration papers [30,31] and LISA [33,49]. They are denoted in the following as BOS [42,50] and LRS [51]. Each has an intersection probability  $p = 1$ , and the largest loops produced are of size  $0.1t$ . The main difference between these models is that, relative to the BOS model, the LRS model has an additional population of very small loops which aim to account for physics at the gravitational backreaction scale. These small loops emit high frequency GWs, and lead to modifications of the SGWB at high frequencies  $f \gg H_0(\Gamma G\mu)^{-1}$  [52]. This can be seen in Fig. 1 which plots the numerical evaluation of Eq. (5) for  $N_c = 2$ ,  $N_k = 0$  for the BOS model (continuous lines) and LRS model (dashed lines) for different values of  $G\mu$ .

Notice that as a consequence, the two models considered give a significantly different spectrum in PTA frequency range only for  $\Gamma G\mu \gtrsim 2 \times 10^{-11}$  (as seen for  $G\mu = 10^{-14}$  and  $\Gamma = 57$  in Fig. 1).

In this work, we consider two cases (i) smooth cosmic string loops with two cusps only (as in Fig. 1), (ii) fixed

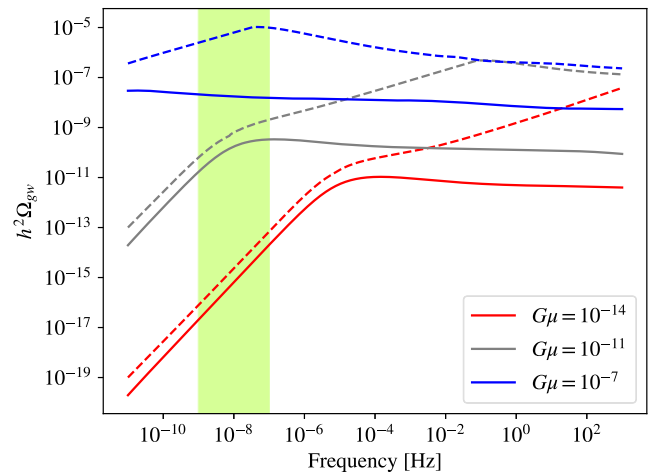


FIG. 1. SGWB from a network of cosmic string loops, expressed in terms of characteristic energy density. The spectrum is computed using the BOS (resp. LRS) loop number density model for the solid (resp. dashed) lines. Here we have taken  $N_c = 2$ ,  $N_k = 0$  leading to  $\Gamma = 57$ . For each model, computations using three different tension values  $G\mu$  are represented. The sensitivity frequency range of EPTA corresponds to the yellow band.



number of cusps,  $N_c = 1$ , but a varying number of kinks  $N_k$  with  $N_k^2/4$  kink-kink collisions. Our aim is to quantify how well the SGWB from cosmic string loops can explain the common red process that is already seen with strong evidence in three PTAs, namely the EPTA, NANOGRV, and PPTA consortiums [1–3].

### III. DATA ANALYSIS METHOD

In this section, we use a Bayesian analysis to perform parameter estimation and model selection. We will not repeat the basis of Bayesian inference in the context of PTA, and instead refer the reader to [1,35] for a more detailed description. In the following subsection, we briefly describe the noise and SGWB models which we test.

We have applied three Bayesian approaches to the six pulsars of the early second EPTA data release (DR2) [1]:

- (1) We directly infer the parameters of each model using the spectral shape of the SGWB as a prior, considering HD correlation between pulsars. This is the most computationally expensive path, which is referred to as “full.”
- (2) Similar to 1, but initially we neglect all correlations between pulsars, assuming an uncorrelated common process. Then, we resample the posterior points using “correct” likelihood *with* HD correlation following the procedure outlined in [38]. This path is significantly faster, we refer to it as “resampling” (RS).
- (3) Finally, we also use the method suggested in [41] and further developed in [40]. There, we first estimate the correlated power at each Fourier bin with a relatively large log-uniform prior: we refer to this as the “free spectrum.” Then, we perform Bayesian analysis using the free spectrum as observational data. This is the computationally cheapest path, which we refer to as the free spectrum (FS) method.

#### A. PTA data analysis principles

In PTA, we work with time residuals, obtained as the difference between observed and predicted time of arrivals of the pulsars’ radio pulses

$$\delta t_i = t_i^{\text{obs}} - t_i^{\text{TM}}(\vec{\beta}_b), \quad (6)$$

where  $i$  is the index of each observation and  $t^{\text{TM}}$  is the timing model attempting to explain the deviations from the observed time-of-arrival using a set of parameters  $\vec{\beta}_b$ . Following the usual path [36,37] we marginalize analytically over timing model parameters.

We assume that the noise in each pulsar consists of the radiometer white noise, intrinsic pulsar spin noise, dispersion measurement and finally scattering variations. These last two components are referred to as chromatic

noises, since they are due to the propagation of the radio pulses through the interstellar medium and depend on the observational radio frequency. We follow the standard procedure [35], initially we infer the noise model in each pulsar individually, and then we use these results as a starting point in the search for the SGWB common to all pulsars.

Besides these individual noise components, we also introduce SGWB into the data model. It is characterized by the spectral shape (subject to its source) and exhibits a particular correlation across pulsar pairs which depends on the angular separation of pulsars in the sky and is described by the Hellings-Downs (HD) curve [34].

We assume that the data does not contain any deterministic signals, with the exception of the two deterministic chromatic signals, referred to as exponential dips, of PSR J1713 + 0747 (see Table I).

We can summarize the data model in the form of the total covariance matrix

$$C_{(ai)(bj)} = \mathcal{N}_{a,(ij)}\delta_{ab} + (C_{a,(ij)}^{\text{RN}} + C_{a,(ij)}^{\text{ChRN}})\delta_{ab} + \Gamma_{ab}C_{(ij)}^{\text{CRN}}, \quad (7)$$

where  $\delta_{(\dots)}$  is the Kronecker delta function,  $a, b$  are indexing the pulsars,  $i, j$  the timing residuals,  $\mathcal{N}_a$  is the white noise covariance matrix of pulsar  $a$ ,  $C_a^{\text{RN}}$  its intrinsic red-noise covariance matrix and  $C_a^{\text{ChRN}}$  its chromatic red-noise covariance matrix.  $C^{\text{CRN}}$  is the common red noise covariance matrix, and  $\Gamma_{ab}$  is the overlap reduction function giving the correlations between pulsars  $a$  and  $b$ . The function  $\Gamma_{ab}$  reduces to  $\delta_{ab}$  for an uncorrelated common red noise and is described by the HD correlations for a SGWB,

$$\Gamma_{ab} = \frac{3}{2}x \ln x - \frac{1}{4}x + \frac{1}{2}(1 + \delta_{ab}), \quad (8)$$

where  $x = \frac{1 - \cos \xi_{ab}}{2}$ ,  $\xi_{ab}$  being the angular separation between pulsars  $a$  and  $b$ .

In PTA data analysis, red stochastic processes are often described via Gaussian processes with truncated Fourier basis (see [37] for details)

$$\delta t_{\text{red}}(t_i) = \sum_{n=1}^{N_f} \left[ a_n \cos\left(\frac{2\pi n t_i}{T}\right) + b_n \sin\left(\frac{2\pi n t_i}{T}\right) \right], \quad (9)$$

where  $T$  is the observational time span and  $a_n, b_n$  are random variables (weights) defined by the spectral shape of the underlying process. In this work, we use  $N_f = 30$  for the common red process while for the individual pulsar noises, we employ a custom number of Fourier bins determined through single pulsar analysis, following the method detailed in [35] (see Table I for an overview of the different  $N_f$  values used for each pulsar noise).

For the chromatic noises, we use a similar approach but introduce a scaling with the observational radio frequency  $\nu$ ,  $\delta t_{\text{red}} \propto \nu^{-\chi}$ , where  $\chi$  is a chromatic index equals to 2 for dispersion measure variations and 4 for scattering variations.

For all intrinsic red noise components attached to each pulsar, we assume a power-law shape for the one-sided power spectral density (PSD) of the residuals

$$S(f; A, \gamma) = \frac{A^2}{12\pi^2} \left( \frac{f}{\text{yr}^{-1}} \right)^{-\gamma} \text{yr}^3, \quad (10)$$

characterized by the spectral index  $\gamma$  and the amplitude  $A$  defined at the reference frequency of 1/year.

Here we assume that the CRN is a SGWB with the one-sided power spectral density given in terms of the fractional energy density by

$$S(f; \vec{\eta}) = \frac{H_0^2}{8\pi^4} \frac{1}{f^5} \Omega_{\text{SGWB}}(f; \vec{\eta}), \quad (11)$$

where  $\vec{\eta}$  are the hyperparameters of the SGWB model considered. The SGWB from a population of SMBHB is well approximated by a power-law model, and takes the same form as in (10), see [10]. Moreover, for the SMBHBs in circular orbits with GW-driven evolution, we expect  $\gamma \approx 13/3$ . The spectral shape for the network of cosmic strings is obtained by integration of Eq. (5) which adds computational time.

In the following analysis, we fix the parameters (and, therefore the level) of the white noise based on each pulsar investigation [35]. We vary the parameters of all red noise components together with the parameters of SGWB. In total, there are  $30 + N_{\text{CRN}}$  parameters, where  $N_{\text{CRN}}$  is the number of hyperparameters of the CRN (varies from model to model).

### B. Full method

This is the standard path one would follow to infer the parameters of each model (pulsars and SGWB). The likelihood for the concatenated array of observations (residuals)  $\vec{\delta t}$  is given by the usual Gaussian form

$$p(\vec{\delta t} | \vec{\eta}) = \frac{1}{\sqrt{|2\pi C|}} \exp \left( -\frac{1}{2} \vec{\delta t}^T C^{-1} \vec{\delta t} \right), \quad (12)$$

where  $C$  is the full covariance matrix of Eq. (7). It is parametrized by  $A_a, \gamma_a$  for each red noise component for each pulsar  $a$  and by parameters describing the SGWB spectrum. We use the Python package ENTERPRISE [53] for its computation (with an extension to include cosmic string models). We sample parameters using PTMCMC sampler [54]. Note that we do not employ the “parallel tempering” features due to technical reasons.

For the SGWB we have used either the power-law model, the BOS or the LRS models with fixed and variable

number of kinks. This is computationally the most expensive (but most rigorous) approach. There are two bottlenecks in the analysis: (i) integration of Eq. (5), and (ii) accounting for the correlation between pulsars (inverting the  $\Gamma_{ab}$  part in the covariance matrix). However, thanks to the low dimensionality of the parameter space of the cosmic string spectrum ( $f, G\mu, N_k$ ), we were able to address point (i) by using numerical interpolation method over a precomputed grid to speed-up the computation of its power spectrum.

We can also perform Bayesian model selection. We use the Bayes factor (BF) as a way to measure which model is preferred by the data. The Bayes factor is equivalent to the posterior odds ratio  $\mathcal{O}$  if we assume (and we do) equal model priors

$$\mathcal{B}_{M_2}^{M_1} = \frac{p(d|M_1)}{p(d|M_2)} = \frac{p(M_1|d) p(M_2)}{p(M_2|d) p(M_1)} = \mathcal{O}_{M_2}^{M_1} \frac{p(M_2)}{p(M_1)}. \quad (13)$$

It is computationally prohibitive to compute the evidence for each model directly using the nested sampling method. Instead, we used the product-space sampling approach described in [55] and already used in several PTA analysis (e.g., in [1,56]). The artificially introduced hyperparameter (model index) allows the sampler also to jump between several models and the BF is given by a ratio of the number of samples accumulated in each model.

### C. Resampling method

This approach is based on importance sampling and described in detail in [38]. The main idea is to perform sampling like in the full method neglecting the correlation between pulsars, taking  $\Gamma_{ab} = \delta_{ab}$  in (7). This way the likelihood is factorized into a product of likelihoods for each pulsar [39]. The likelihood computational time is reduced by a factor 6 for six pulsars, but this gain increases as the number of pulsars analyzed increases. Once it is done, we resample (reweigh) the posteriors by recomputing the likelihood for each sample using  $\Gamma_{ab}$  defined by the HD curve. As a result, it is possible to parallelize the calculation and therefore gain even more speed. This method works well if the difference in the posteriors obtained with and without HD correlation is not very large (not disjoint distributions).

We can use the same approach to get a cheap estimation of the Bayes factors. To the best of our knowledge, it was not used before, so we provide here some detailed computation.

We refer the reader to Sec. II of [38] for details on the resampling of the posterior from an approximate model “ $A$ ” to a target posterior “ $T$ .” We apply this method to the product space approach for computing the Bayes factor. We consider two hypermodels  $\mathcal{M}, \mathcal{M}^{\text{HD}}$  with the same model parameters  $\theta$  and hyperparameter  $n$  indexing the models describing the common red noise we want to compare. The difference between the two hypermodels is in the likelihood

computation: we consider uncorrelated common process in the hypermodel  $\mathcal{M}$ . Computations of Bayes factor with  $\mathcal{M}$  are significantly faster and more reliable, we want to evaluate the Bayes factor of models in  $\mathcal{M}^{\text{HD}}$  by resampling the hyperchain (product space)  $\mathcal{M}$ . We repeat the steps outlined in Sec. 3 of [55] but introducing the approximate likelihood in the posterior probability of  $n$  for  $\mathcal{M}^{\text{HD}}$

$$\begin{aligned} P(n|\mathcal{D}, \mathcal{M}^{\text{HD}}) &= \frac{\pi(n)}{Z_{\mathcal{M}^{\text{HD}}}} \int \mathcal{L}^{\text{HD}}(\theta_n) \pi(\theta_n|n) d\theta_n \\ &= \frac{\pi(n)}{Z_{\mathcal{M}^{\text{HD}}}} \int \frac{\mathcal{L}^{\text{HD}}(\theta_n)}{\mathcal{L}(\theta_n)} \mathcal{L}(\theta_n) \pi(\theta_n|n) d\theta_n \\ &\approx \frac{\pi(n)}{Z_{\mathcal{M}^{\text{HD}}}} \frac{1}{N_s^{(n)}} \sum_{i=1}^{N_s^{(n)}} w_i^{(n)} Z_n. \end{aligned} \quad (14)$$

Here we introduced  $\pi(n)$  the prior on the model index (chosen to be uniform),  $\mathcal{L}^{\text{HD}}/\mathcal{L}$  are likelihood functions with/without HD spatial correlations and  $Z_{\mathcal{M}^{\text{HD}}}$  is the evidence of the hypermodel with HD (in our case just normalization factor which cancels out). In the last line, we have approximated the integral using Monte Carlo approach, with  $N_s^{(n)}$  being the number of samples corresponding to the model associated to  $n$  in the hypermodel  $\mathcal{M}$  posterior. The weights  $w_i^{(n)} = \mathcal{L}^{\text{HD}}(\theta_n^{(i)})/\mathcal{L}(\theta_n^{(i)})$  are computed at each posterior points (indexed by  $i$ ) of  $\mathcal{M}$ , the  $\theta_n$  being the  $\theta$  parameters attached to the model  $n$ .

The Bayes factor [assuming uniform prior across the models  $\pi(n)$ ] between the models within each hypermodel posterior is given as

$$B_{M_b}^{M_a} \equiv B_{a/b} = \frac{P(\lfloor n \rfloor = n_a | \mathcal{D}, \mathcal{M})}{P(\lfloor n \rfloor = n_b | \mathcal{D}, \mathcal{M})}, \quad (15)$$

where  $\lfloor \cdot \rfloor$  is the floor function.

Applying this expression to the resampled hypermodel posterior, we obtain

$$B_{a/b}^{\mathcal{M}^{\text{HD}}} = \frac{Z_a^{\text{HD}}}{Z_b^{\text{HD}}} = \frac{\bar{w}_{(a)}}{\bar{w}_{(b)}} \times B_{a/b}^{\mathcal{M}}, \quad (16)$$

where  $\bar{w}_{(n)} = \frac{1}{N_s^{(n)}} \sum_{i=1}^{N_s^{(n)}} w_i^{(n)}$  is a posterior average weight of the model indexed by  $n$ . As previously mentioned, the Bayes factor for the hypermodel  $\mathcal{M}$  (without HD correlations) is simply given by  $B_{a/b}^{\mathcal{M}} \approx N_s^{(n_a)}/N_s^{(n_b)}$ . The performance of this method depends on the sampling efficiency, as described in [38]. Note that the novelty of this method is in the efficient evaluation of Bayes factor between *two or more* GWB models via resampling.

## D. Free spectrum method

As mentioned, this method was suggested in [41] and further developed in [40]. The main idea is to get rid of the high dimensionality of the problem by marginalizing over the single pulsar noise parameters. We want to estimate the PSD assuming the HD correlation at each frequency bin  $f_k = k/T$ , where  $T$  is the total time span of the array. Using MCMC we get a posterior distribution for the amplitude  $\rho_k$  at each frequency bin, where

$$\rho_k^2 = \frac{S(f_k)}{T}, \quad (17)$$

where  $S(f)$  is the one-sided PSD of the SGWB we want to characterize, see Eq. (11). We assume that all bins are independent and impose log-uniform prior for each amplitude,  $\log_{10} \rho_k \in [-12, -4]$ , a justification for this range is given in Sec. IV B.

We use a kernel density estimator to obtain a smooth probability distribution function  $p(\rho_k|\vec{\delta}t)$ . In this study, we use Gaussian kernels, and their bandwidth is selected using Scott's rule of thumb [57]. We use the estimated PSD of a common process with HD correlation as our new data. We can then build a factorized likelihood that depends only on the parameters of the SGWB  $\vec{\eta}$ , following [40], we get

$$\begin{aligned} p(\vec{\delta}t|\vec{\eta}) &\approx \prod_{k=1}^{N_f} \int d\rho_k p(\vec{\delta}t|\rho_k) p(\rho_k|\vec{\eta}) \\ &\propto \prod_{k=1}^{N_f} \int d\rho_k \frac{p(\rho_k|\vec{\delta}t)}{p(\rho_k)} p(\rho_k|\vec{\eta}). \end{aligned} \quad (18)$$

We have used the following notation:  $p(\rho_k|\vec{\delta}t)$  are the posterior distributions of the  $\rho_k$  describing our observations in Fourier domain,  $p(\rho_k)$  is the prior on the  $\rho_k$  and  $p(\rho_k|\vec{\eta})$  is the probability to have  $\rho_k$  given the CRN parameters  $\vec{\eta}$ . For the latter, we consider that the PSD is perfectly deterministic such that

$$p(\rho_k|\vec{\eta}) = \delta(\rho_k - \rho_k^{\text{SGWB}}(\vec{\eta})), \quad (19)$$

where  $\rho_k^{\text{SGWB}}(\vec{\eta})$  is derived from the PSD of a corresponding SGWB model computed at  $f_k = k/T$  for a set of background parameters  $\vec{\eta}$ .

Note that the Fourier frequency bins are not entirely independent as the PTA data are not evenly sampled and pulsars do not have the same observational duration. We have computed the associated Pearson correlation coefficient matrix  $P_{i,j} = C_{i,j}/\sqrt{C_{i,i}C_{j,j}}$  where  $C_{i,j}$  is the covariance matrix of the  $\rho$ 's. We have found that the frequency bins are approximately independent: the average of the nondiagonal coefficients being  $\langle |P_{i,j}| \rangle_{i < j} \approx 0.02$ .

Taking all into account, the likelihood of Eq. (18) takes the form

$$p(\vec{\delta t}|\vec{\eta}) \propto \prod_{k=1}^{N_c} p(\rho_k^{\text{SGWB}}(\vec{\eta})|\vec{\delta t}). \quad (20)$$

This method is extremely fast and drastically reduces the dimensionality of the problem, making it possible to use the Nested sampling [58,59] algorithm to obtain the evidence for each model and posterior distributions of the SGWB parameters  $\vec{\eta}$ . We used a particular implementation of nested sampling, DYNesty, described in [60,61].

#### IV. RESULTS

Before discussing our results, we should mention that the six pulsars dataset considered here is noninformative about the presence of the HD correlations, as was shown in [1,35]. Consequently, we expect that the results of the analysis will be similar (though not identical) using CURN or GWB assumptions for the CRN. The main purpose of this paper is to introduce the methodology which will be applied to the extended (25-pulsar) EPTA dataset.

We start by considering the cosmic string models with the fixed average number of cosmic string bursts per oscillation to  $N_c = 2$  and  $N_k = N_{kk} = 0$ , leading to  $\Gamma = 57$ , which is expected by a population of smooth loops [62]. Later we extend our analysis by allowing the number of kinks to vary (and so the value of  $\Gamma$ ).

#### A. Parameter estimation

##### 1. Smooth loops

When we infer a SGWB from a network of smooth loops with  $N_k = 0$  and  $N_c = 2$  (alongside the individual pulsar noise models), we find a very constrained distribution for the tension  $G\mu$  for both loop density distribution models, see Fig. 2. As one can see the posterior distributions given by the full, the RS, and the FS methods agree very well, though we find that the FS method gives slightly broader posteriors. The string tension 90% credible (symmetric) intervals for each Bayesian method can be found in the two first lines of Table II. These intervals show nice consistency across all three Bayesian methods:  $\log_{10} G\mu \sim -10.1$  (resp.  $-10.6$ ) for the BOS (resp. LRS) model.

These results can be understood as follows. For the BOS model, the SGWB exhibits a peak around  $f \sim F_0 \equiv 3H_0(\Gamma G\mu)^{-1}$  before decreasing to reach a plateau at higher frequencies (see Fig. 1) [52]. According to the posterior for  $G\mu$ , the peak is in the middle of the PTA band, thus the spectral index of  $\Omega_{\text{SGWB}}$  transits from 1 to 0 (flat spectrum) throughout the PTA frequency range. This is comparable to the value of  $2/3$  (corresponding to  $\gamma = 13/3$ ) expected for a SMBHB background. As can be seen in Fig. 5, the BOS model slope is very compatible with the spectral estimated from the data.

Regarding the LRS model, the extra population of small loops (see [52]) is responsible for the fact that instead of decreasing to a plateau for  $f > F_0$ , the characteristic energy

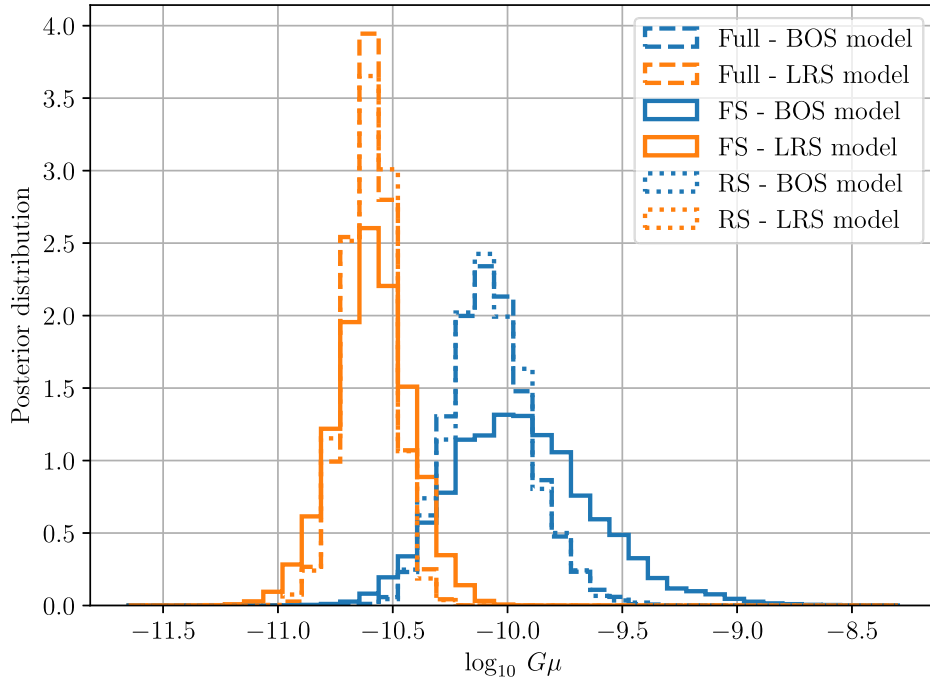


FIG. 2. Comparison of the string tension posteriors (for two string models, BOS and LRS) obtained with the full method (dashed lines), resampling (RS) method (dotted lines), and with the free spectrum (FS) method (solid lines). We assume here that the loops are populated by two cusps, leading to  $\Gamma = 57$ .



TABLE II. For each Bayesian analysis method (columns), we write the cosmic string parameters posteriors for each loop distribution (BOS/LRS) with a varying number of kinks (\_kk) or without. The 5% and 95% quantiles are used to set up the credible interval. The second column quotes the linear Bayes factors comparing each of the cosmic string (CS) models against a power-law (PL) PSD for the SGWB.

	Full method		RS method		FS method	
	CS posterior	BF (PL/CS)	CS posterior	BF (PL/CS)	CS posterior	BF (PL/CS)
BOS	$\log_{10} G\mu = -10.08^{+0.32}_{-0.26}$	0.3	$-10.08^{+0.30}_{-0.27}$	0.2	$-9.95^{+0.56}_{-0.45}$	0.1
LRS	$\log_{10} G\mu = -10.60^{+0.17}_{-0.17}$	0.2	$-10.60^{+0.16}_{-0.18}$	0.3	$-10.59^{+0.27}_{-0.28}$	0.2
BOS_kk	$\log_{10} G\mu = -8.92^{+1.12}_{-1.42}$	0.2	$-8.98^{+1.17}_{-1.38}$	0.3	$-8.66^{+1.24}_{-1.65}$	0.1
	$N_k = 105^{+86}_{-96}$		$101^{+90}_{-96}$		$100^{+88}_{-87}$	
LRS_kk	$\log_{10} G\mu = -10.90^{+0.29}_{-0.27}$	0.2	$-10.89^{+0.29}_{-0.27}$	0.3	$-10.86^{+0.35}_{-0.40}$	0.2
	$N_k = 102^{+89}_{-92}$		$103^{+89}_{-95}$		$104^{+85}_{-92}$	

density rather increases again up to a second peak at very high frequencies compared to the PTA band, see Fig. 1. The fractional energy density follows there a power-law of spectral index  $\approx 0.4$ , which also fits well the spectral shape of the data (see Fig. 5).

As an alternative, we also consider a two-component SGWB model: a SGWB generated by a population of circular binary black holes with  $\gamma = 13/3$  plus a stochastic signal from the cosmic strings.

We find a very strong correlation between those two components as one can see in Fig. 3. The 2D-posteriors exhibit a pronounced L-shape, meaning that both model can

explain the data. Based on the obtained posterior distribution, we can extract an upper limit which we evaluate to be  $\log_{10} G\mu < -9.9$  (resp.  $-10.5$ ) for the BOS (resp. LRS) model using the full method, and it corresponds to 95% confidence. In addition, the evaluation obtained with the two fast methods (RS, [FS]) give very consistent upper bounds:  $\log_{10} G\mu < -9.8[-9.5]$  (resp.  $-10.5[-10.4]$ ) for the BOS (resp. LRS model). The slightly broader posterior obtained with FS earlier leads to a less stringent upper bound for the BOS model. One can notice that we constrain better the tension,  $G\mu$ , for the LRS model, this is due to the excess of power that the small loop population introduces.

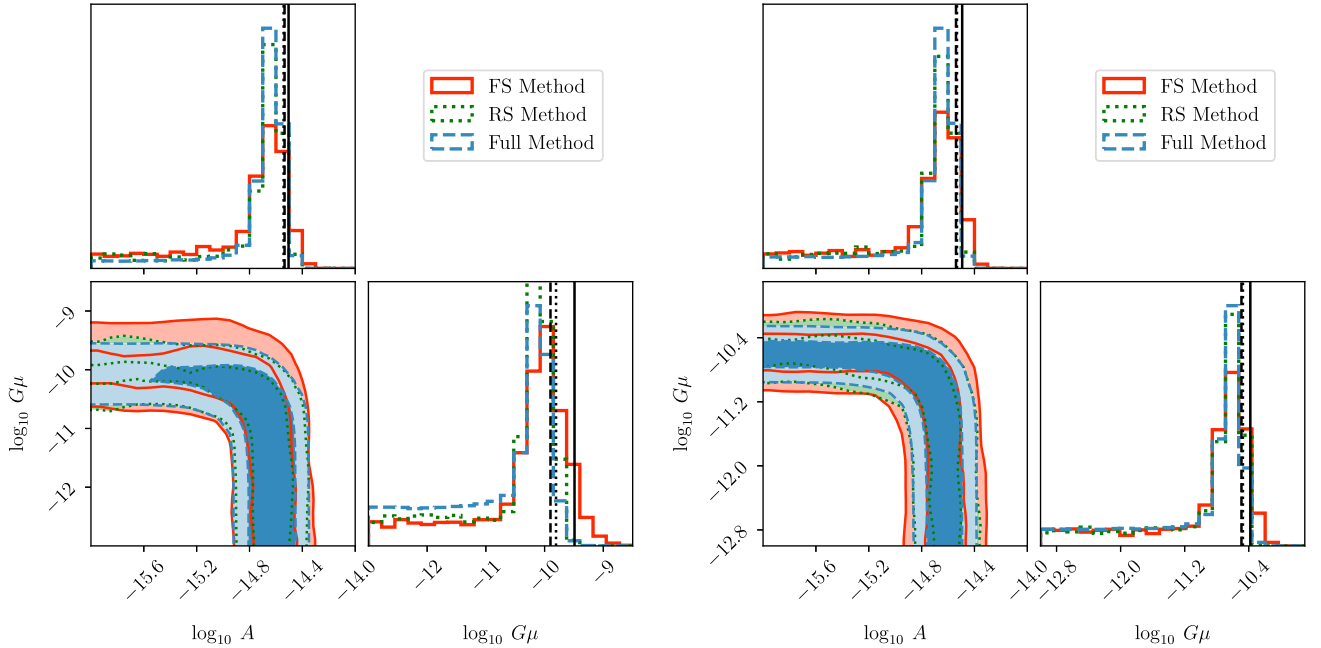


FIG. 3. Left panel: posterior for the two-component SGWB model composed of (i) a signal originating from a population of circular GW-driven SMBHB, parametrized by its PSD amplitude  $\log_{10} A$  at  $f = 1/\text{year}$ , and (ii) a SGWB from smooth CS loops background using the BOS loop number density model; different lines styles (dashed, dotted, solid) corresponding to three methods (full, RS, FS) show good consistency. The 95-quantile for each SGWB parameter posterior is plotted in black, using the line style associated with its respective method. Right panel: the same for the LRS loop number density model.

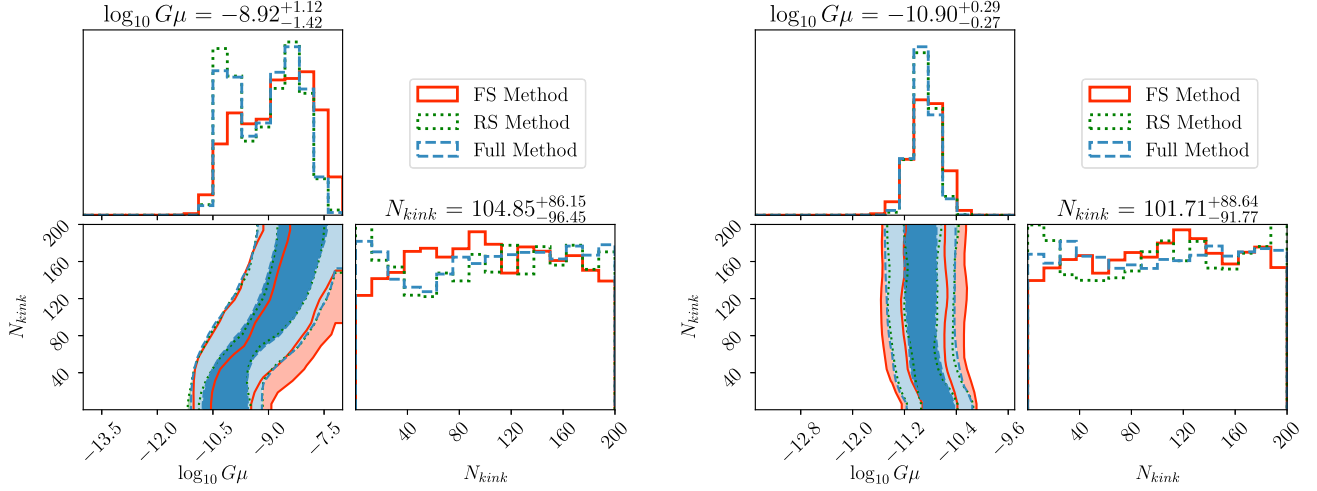


FIG. 4. Left panel: posterior for two-dimensional (string tension and the average number of kinks on loops of the network) BOS model; different lines styles corresponding to three methods (full, RS, FS) show full consistency. Right panel: the same for the LRS loop number density model.

## 2. Kinky loops

Next we also vary the number of kinks (and thus kink-kink collisions) in addition to the string tension, considering two-dimensional models  $(\log_{10} G\mu, N_k)$ . The PSD of the SGWB created by such a model is calculated from Eq. (5) by setting  $N_c = 1$  (in order to still have some GW power at the low number of kinks) and  $N_{kk} = N_k^2/4$ .

We took a uniform prior for  $N_k$  letting it vary between 0 and 200, the same prior was used in the LVK analysis [31]. This large prior on  $N_k$  accounts for theoretical uncertainties on the initial number of kinks at loop creation and on the efficiency of the gravitational backreaction that is expected to smooth out the loop.

In the case of the BOS model, if we increase the number of kinks (and so  $\Gamma$ ), the SGWB in the PTA band now corresponds to the transition between the peak of the spectrum and the high frequency plateau and thus increases the characteristic slope of its PSD while decreasing the amplitude. This increase in spectral index will be disfavored as it will be too steep to fit the data properly. However, to compensate the decrease in amplitude, an increase of the string tension  $G\mu$  can both correct the amplitude and the spectral index by placing the high frequency plateau in the PTA frequency band. One can see such interplay between  $G\mu$  and  $N_k$  in the left panel of Fig. 4. The data equally allows two joined solutions: low tension with low number of kinks and very kinky loops ( $N_k \gtrsim 120$ ) with high tension,  $\log_{10} G\mu \sim -8.3$ .

Contrary to the BOS case, the PSD slope of the LRS increases only slightly at low frequency with increasing  $N_k$ . Once  $F_0 \lesssim 1/T \sim 10^{-9}$ , the spectrum is dominated by the extra population of small loops that the LRS model introduces and the associated power is now independent of  $\Gamma$  and proportional to  $G\mu$  [32]. Therefore, the only shift in amplitude is caused by the domination of kinks (instead

of cusp) in the GW power emission leading to an increase by a factor  $\sim 1.6$  [32], that can be simply compensated by a slight decrease in the tension leading to  $\log_{10} G\mu \sim 10.9$ . As a result, adding kinks has practically no effect in the LRS model, and we recover the prior for  $N_k$ , as seen the right panel of Fig. 4.

According to the posterior, the model with a high number of kinks is supported by the observed data, especially for BOS where it suggests a possible very high number of kinks. For example,  $N_k = 120$  corresponds to  $\Gamma \approx 960$ , which is well above the value of 50 expected in the latest simulations of cosmic string loops. Finally, all three methods (depicted with different line styles in Fig. 4) show very consistent results, even if the FS method struggles more to recover the double peak solution in the BOS model case.

## B. Model comparison

The fact that the cosmic string parameters are well constrained does not tell us how well it explains the observed data. The Bayes factor, defined as the ratio of evidence between two models, is often used to quantify the ability of two different models to fit a set of data. Note that the Bayes factor can tell us which model is preferred based on the observations, but still does not tell us how well it describes the data. In what follows, we want to compute the Bayes factors of the considered cosmic string models against the power-law (PL) (SMBHB-inspired) model. Results are reported in Table II. We find that the cosmic string models are slightly preferred over PL model thanks to its simplicity: only one parameter. However, Fig. 5 shows that all models fit the spectral shape similarly well.

Considering the two-component model of SGWB (SMBHB + CS, see previous subsection), we have found no statistical support for this complex model as compared

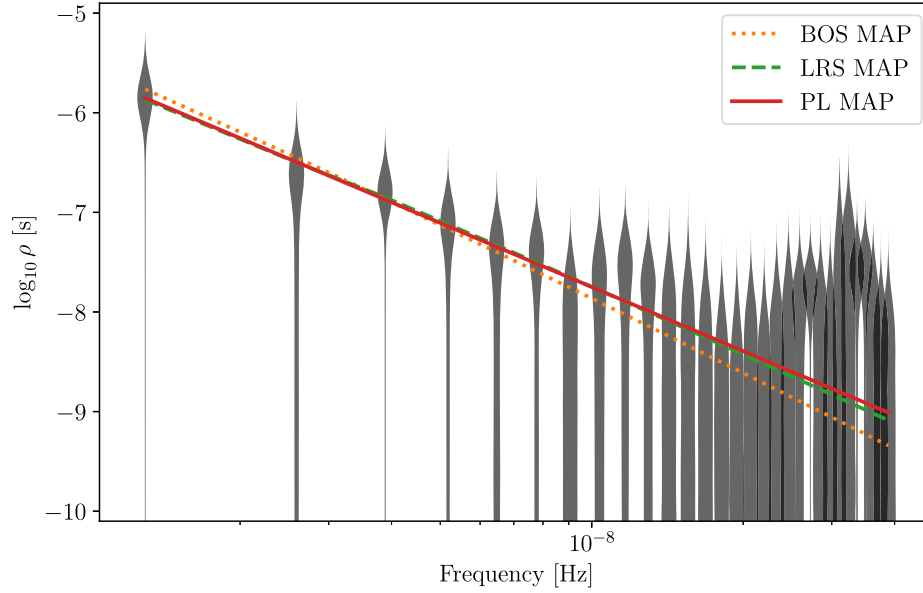


FIG. 5. Posterior distributions of the 30  $\rho_k$  coefficients (in  $\log_{10}$ -scale). We over-plotted the best fit [using the FS likelihood of Eq. (20)] for three different PSDs: powerlaw, SGWB using BOS and LRS models (in the case of smoothed loops). We see that all three spectra behave in a similar way at low frequency bins.

to a simple SMBHB,  $\mathcal{B}_{\text{SMBHB}}^{\text{SMBHB}+\text{CS}} \approx 1$  for both BOS and LRS cosmic string networks (using full, RS and FS methods).

We have demonstrated earlier that the kinks had no strong impact and were not constrained by the data in the models with the variable number of kinks. This is also reflected in the evaluated Bayes factors.

Both full and the RS methods give very consistent Bayes factors for all models, as shown in Table II.

The Bayes factors computed with FS method depend on the prior chosen for evaluating the FS amplitude components  $\rho_k$ . The main reason is that power at all low frequency bins is not fully constrained. The posteriors have a long tail toward the low values of power (see Fig. 5). Choosing a wider prior makes those tails longer, spreading the probability across a long range of  $\rho_k$  (draining it from the “bumps” located at high power). Despite the fact that the tails are thin, they go all the way to minus infinity in log-scale (zero in power) and the probability distribution is then strongly affected (through normalization). So we are facing the question of what we should use as a prior range for the  $\log_{10} \rho_k$ ?

The choice of the prior range has a direct impact on the evidence calculation: some values of the SGWB model parameters have no support (zero contribution to the evidence integral) because of the truncation of the  $\rho_k$  imposed by the prior (which is not the case in the full or RS approach). Even more important is the indirect impact caused by the correlation of pulsar spin noise (red noise intrinsic to each pulsar) and common red noise at the lowest frequencies. In Fig. 6, we show the inferred distribution for the parameters of the spin red noise in

one of the best pulsars, PSR J1909-3744. Extending the prior range for  $\rho_k$  leads to steeper and stronger spin red noise in that pulsar (dashed brown line compared to the solid brown line). This implies that the (artificially)

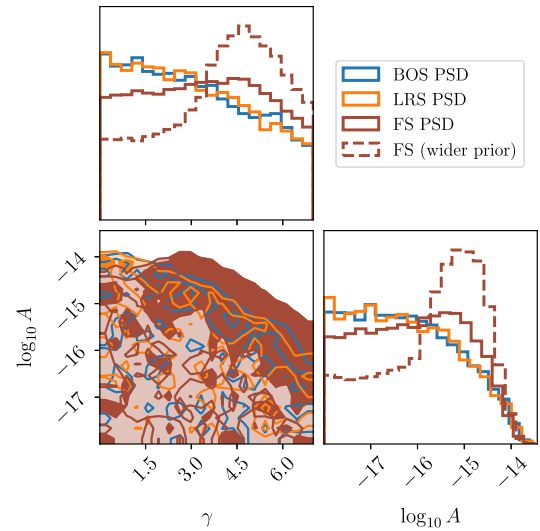


FIG. 6. Marginalized red noise parameters posteriors for the pulsar J1909-3744 obtained with the full method including HD correlation for the common red noise process. Brown dashed and solid lines were obtained using different prior on the power in the Fourier bins (passing from  $-10$  to  $-15$  for the  $\log_{10} \rho_k$  lower bound). The solid line (restricted prior) suggests a truncated correlation of the common red noise with the spin red noise of J1909-3744. We also plotted for comparison those posteriors when using as common red noise, a SGWB from cosmic string following the BOS (blue line)/LRS (orange line) models.

overconstrained prior on  $\rho_k$  truncates the noise estimation in the pulsar J1909-3744. We empirically determined a fiducial value of power in terms of  $\log_{10}\rho_{\text{th}}$  (that can be roughly regarded as common to all frequency bins) which represents the threshold below which our data becomes insensitive. We selected a symmetric range centered around this specific value, which we set at  $\log_{10}\rho_{\text{th}} = -8$ . For the upper limit, we opted for  $-4$ , leading to a lower bound of  $-12$  for the  $\log_{10}\rho$  prior. This prior answers direct (albeit partially) and indirect (by preserving correlations) impact on the Bayes factor while avoiding an overflow of the prior range.

In addition, it is worth noting<sup>5</sup> that the posteriors given by the FS method typically have a broader spread compared to other methods. Moreover, the FS method encounters difficulties in accurately capturing the double peak feature observed in the  $G\mu$  posterior for the BOS model in Fig. 4. This issue is mainly due to the width of the prior for the  $\rho_k$  parameters. Indeed, if a  $\rho_k$  posterior is not well-constrained, which is the case for several frequency bins as evident in Figure 5, taking a wider prior results in a longer tail and less constrained posterior at its upper bound (due to normalization). This gives a likelihood function [Eq. (20)] with less pronounced peaks, leading to the spread in the posterior distributions. By running the FS method using a narrower prior range of  $[-10, -4]$  for the  $\rho_k$ , we recover the posterior distribution which is much closer to the one obtained with the full method.

In general, we have found the choice for the prior for the free spectrum evaluation (or rather a lack of rigorously defined cut) is a weak point of the FS approach in computing the evidence, despite that it perfectly falls into Bayesian philosophy.

Accumulating high-quality data and using more pulsars in the array, will hopefully result in the well-constrained power at low frequencies. Consequently, this will make the choice of the prior for the  $\rho_k$  parameters much less relevant for evaluation of the Bayes factors using the FS method.

## V. CONCLUSION AND DISCUSSION

In this work, we demonstrated that the free spectrum and resampling methods are very powerful tools to determine rapidly the posterior distributions for the parameters describing the SGWB. Moreover, those methods (taking into account caveats for FS method discussed above) could be used to evaluate the Bayes factor between various SGWB models. We have demonstrated these methods using the 6-pulsar early DR2 EPTA dataset and comparing several models of SGWB produced by a network of cosmic strings.

For this 6-pulsars only dataset, we could perform rigorous inference (full method) of SGWB parameters (in addition to pulsar noise parameters) and Bayes factors evaluation Table II. Note that the full method took several days to complete, whereas the execution time for RS/FS methods is reduced to hours/minutes. Extending the data by including more pulsars, more back-ends and new observations will hugely explode the dimensionality of the problem and make the full method computationally not tractable. The computational power scaling quadratically with the number of pulsars and the number of observations. We hope that this work convinces the reader of the validity of the fast and approximate methods which have to be used in the future PTA data analysis. The attractiveness of FS method is that it is not very sensitive to the number of pulsars in the array. Of course the evaluation of the free spectrum does heavily depend on the volume of the data and dimensionality of parameters space (pulsars noise models), but it can then be cheaply used to infer parameters for multiple SGWB models. The resampling method is also relatively cheap, the sampling step scales linearly with number of pulsars in the array and can be efficiently parallelized (if needed). Even though we have demonstrated the fast methods using simplistic power-law model for SMBHB and two cosmic string SGWB models, the methods are generic and can be used to infer parameters of any SGWB.

The second result of this paper are the constraints on the cosmic string tension  $G\mu$  for the BOS and LRS models. Assuming that the observed red noise process is GWs, we have obtained the preferred value for the string tension  $G\mu \approx 10^{-10.1}$  (resp.  $10^{-10.6}$ ) for BOS (resp. LRS). Moreover, we find that the power-law model with HD correlations is slightly disfavored compared to the simpler one-parameter cosmic strings model. The posterior is not informative about the number of kinks when we consider the two-parameter  $(G\mu, N_k)$  cosmic string models. However, it also implies that the possibility of having a large number of kinks is not ruled out, though it might be disfavored on the theoretical grounds. Considering a two-component SGWB model (SMBHB + CS) shows a strong correlation between two components, and we set the upper bounds  $\log_{10} G\mu \lesssim -9.9$  (BOS model) and  $\log_{10} G\mu \lesssim -10.5$  at 95% (LRS model) confidence.

Compared to previous EPTA constraints [63,64] on  $\log_{10} G\mu$  which were obtained using the EPTA sensitivity curve, here we use more up to date models and consider directly the PSD of the cosmic string SGWB through a wide range of frequencies to obtain our constraints. Our upper bounds on the string tension are well under the one given by CMB experiment  $\log_{10} G\mu \lesssim -7$  [65] and the LVK collaboration [31] for the BOS model ( $\log_{10} G\mu \lesssim -8$ ). However, for the LRS model, due to the population of small loops emitting at higher frequencies, the LVK constraint is more stringent,  $\log_{10} G\mu \lesssim -14$ .

<sup>5</sup>We thank the referee for pointing us in this direction.



This paper lays the path for the search and interpretation of a SGWB signal of any origin. The fast methods suggested here will be applied to the extended EPTA dataset containing 25 pulsars [5]. In addition to EPTA data, the new dataset which combines the most sensitive observations from all PTAs, International PTA data combination, is being produced. We expect the future data to be more sensitive to GWs and give better constraints on the string tension in BOS and LRS models.

## ACKNOWLEDGMENTS

H. Q.-L., P. A., and D. S. would like to thank the University of Geneva for hospitality while this work was in progress. We are grateful to Chiara Caprini for very useful discussions. H. Q.-L. thanks Institut Polytechnique de Paris for funding his PhD. The work of P. A. is supported by the Wallonia-Brussels Federation Grant ARC No 19/24-103. S. B. and H. Q.-L. acknowledge support from ANR-21-CE31-0026, project MBH\_waves. D. S. is grateful to CERN for hospitality.

- 
- [1] S. Chen *et al.*, *Mon. Not. R. Astron. Soc.* **508**, 4970 (2021).
  - [2] N. Arzoumanian *et al.*, *Astrophys. J. Lett.* **905**, L34 (2020).
  - [3] B. Goncharov *et al.*, *Astrophys. J. Lett.* **917**, L19 (2021).
  - [4] J. Antoniadis *et al.*, *Mon. Not. R. Astron. Soc.* **510**, 4873 (2022).
  - [5] J. Antoniadis *et al.* (EPTA Collaboration), *Astron. Astrophys.* **678**, A50 (2023).
  - [6] G. Agazie *et al.*, *Astrophys. J. Lett.* **951**, L8 (2023).
  - [7] D. J. Reardon *et al.*, *Astrophys. J. Lett.* **951**, L6 (2023).
  - [8] H. Xu *et al.*, *Res. Astron. Astrophys.* **23**, 075024 (2023).
  - [9] G. Agazie *et al.* (NANOGrav Collaboration), *Astrophys. J. Lett.* **952**, L37 (2023).
  - [10] A. Sesana, *Mon. Not. R. Astron. Soc.* **433**, L1 (2013).
  - [11] E. S. Phinney, [arXiv:astro-ph/0108028](https://arxiv.org/abs/astro-ph/0108028).
  - [12] A. Sesana, *Classical Quantum Gravity* **30**, 244009 (2013).
  - [13] J. Ellis and M. Lewicki, *Phys. Rev. Lett.* **126**, 041304 (2021).
  - [14] S. Blasi, V. Brdar, and K. Schmitz, *Phys. Rev. Lett.* **126**, 041305 (2021).
  - [15] L. Bian, J. Shu, B. Wang, Q. Yuan, and J. Zong, *Phys. Rev. D* **106**, L101301 (2022).
  - [16] Z.-C. Chen, Y.-M. Wu, and Q.-G. Huang, *Astrophys. J.* **936**, 20 (2022).
  - [17] M. Hindmarsh and J. Kume, *J. Cosmol. Astropart. Phys.* **04** (2023) 045.
  - [18] H. B. Nielsen and P. Olesen, *Nucl. Phys.* **B61**, 45 (1973).
  - [19] T. W. B. Kibble, *J. Phys. A* **9**, 1387 (1976).
  - [20] T. W. B. Kibble, *Nucl. Phys.* **B252**, 227 (1985); **B261**, 750 (E) (1985).
  - [21] R. Jeannerot, J. Rocher, and M. Sakellariadou, *Phys. Rev. D* **68**, 103514 (2003).
  - [22] C. Ringeval, M. Sakellariadou, and F. Bouchet, *J. Cosmol. Astropart. Phys.* **02** (2007) 023.
  - [23] C. J. A. P. Martins and E. P. S. Shellard, *Phys. Rev. D* **73**, 043515 (2006).
  - [24] K. D. Olum and V. Vanchurin, *Phys. Rev. D* **75**, 063521 (2007).
  - [25] M. Hindmarsh, S. Stuckey, and N. Bevis, *Phys. Rev. D* **79**, 123504 (2009).
  - [26] D. Matsunami, L. Pogosian, A. Saurabh, and T. Vachaspati, *Phys. Rev. Lett.* **122**, 201301 (2019).
  - [27] M. Hindmarsh, J. Lizarraga, A. Urrio, and J. Urrestilla, *Phys. Rev. D* **104**, 043519 (2021).
  - [28] P. Auclair, D. A. Steer, and T. Vachaspati, *Phys. Rev. D* **101**, 083511 (2020).
  - [29] P. Auclair, K. Leyde, and D. A. Steer, *J. Cosmol. Astropart. Phys.* **04** (2023) 005.
  - [30] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. D* **97**, 102002 (2018).
  - [31] R. Abbott *et al.* (LIGO Scientific, Virgo, and KAGRA Collaborations), *Phys. Rev. Lett.* **126**, 241102 (2021).
  - [32] P. Auclair, J. J. Blanco-Pillado, D. G. Figueroa, A. C. Jenkins, M. Lewicki, M. Sakellariadou, S. Sanidas, L. Sousa, D. A. Steer, J. M. Wachter, and S. Kuroyanagi, *J. Cosmol. Astropart. Phys.* **04** (2020) 034.
  - [33] P. Auclair *et al.* (LISA Cosmology Working Group), *Living Rev. Relativity* **26**, 5 (2023).
  - [34] R. W. Hellings and G. S. Downs, *Astrophys. J.* **265** (1983).
  - [35] A. Chalumeau *et al.*, *Mon. Not. R. Astron. Soc.* **509**, 5538 (2021).
  - [36] R. van Haasteren and Y. Levin, *Mon. Not. R. Astron. Soc.* **428**, 1147 (2012).
  - [37] R. van Haasteren and M. Vallisneri, *Phys. Rev. D* **90**, 104012 (2014).
  - [38] S. Hourihane, P. Meyers, A. Johnson, K. Chatziioannou, and M. Vallisneri, *Phys. Rev. D* **107**, 084045 (2023).
  - [39] S. R. Taylor, J. Simon, L. Schult, N. Pol, and W. G. Lamb, *Phys. Rev. D* **105**, 084049 (2022).
  - [40] W. G. Lamb, S. R. Taylor, and R. van Haasteren, *Phys. Rev. D* **108**, 103019 (2023).
  - [41] C. J. Moore and A. Vecchio, *Nat. Astron.* **5**, 1268 (2021).
  - [42] J. J. Blanco-Pillado, K. D. Olum, and B. Shlaer, *Phys. Rev. D* **89**, 023512 (2014).
  - [43] T. Damour and A. Vilenkin, *Phys. Rev. D* **64**, 064008 (2001).
  - [44] C. Ringeval and T. Suyama, *J. Cosmol. Astropart. Phys.* **12** (2017) 027.
  - [45] P. Binetruy, A. Bohe, T. Hertog, and D. A. Steer, *Phys. Rev. D* **82**, 083524 (2010).
  - [46] P. Binetruy, A. Bohe, T. Hertog, and D. A. Steer, *Phys. Rev. D* **82**, 126007 (2010).
  - [47] A. Vilenkin and E. P. S. Shellard, *Cosmic Strings and Other Topological Defects*, Cambridge Monographs on

- Mathematical Physics (Cambridge University Press, Cambridge, England, 1994), <https://books.google.fr/books?id=alpJQgAACAAJ>.
- [48] N. Aghanim *et al.* (Planck Collaboration), *Astron. Astrophys.* **641**, A6 (2020); **652**, C4(E) (2021).
  - [49] P. Auclair *et al.*, *J. Cosmol. Astropart. Phys.* **04** (2020) 034.
  - [50] J. J. Blanco-Pillado and K. D. Olum, *Phys. Rev. D* **96**, 104046 (2017).
  - [51] L. Lorenz, C. Ringeval, and M. Sakellariadou, *J. Cosmol. Astropart. Phys.* **10** (2010) 003.
  - [52] P. G. Auclair, *J. Cosmol. Astropart. Phys.* **11** (2020) 050.
  - [53] J. A. Ellis, M. Vallisneri, S. R. Taylor, and P. T. Baker, ENTERPRISE: Enhanced Numerical Toolbox Enabling a Robust Pulsar Inference Suite, Zenodo v3.0.0 (2020), <https://doi.org/10.5281/zenodo.4059815>.
  - [54] J. Ellis and R. van Haasteren, jellis18/ptmcmcsampler: Official release, 10.5281/zenodo.1037579 (2017).
  - [55] S. Hee, W. J. Handley, M. P. Hobson, and A. N. Lasenby, *Mon. Not. R. Astron. Soc.* **455**, 2461 (2015).
  - [56] Z. Arzoumanian *et al.*, *Astrophys. J.* **859**, 47 (2018).
  - [57] D. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*, Wiley Series in Probability and Statistics (John Wiley & Sons, New York, 2015), <https://books.google.fr/books?id=XZ03BwAAQBAJ>.
  - [58] J. Skilling, in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, American Institute of Physics Conference Series, Vol. 735, edited by R. Fischer, R. Preuss, and U. V. Toussaint (2004), pp. 395–405, 10.1063/1.1835238.
  - [59] J. Skilling, *Bayesian Anal.* **1**, 833 (2006).
  - [60] S. Koposov *et al.*, JOSHSPEAGLE/DYNesty: v2.1.1 (2023), <https://doi.org/10.5281/zenodo.8408702>; J. S. Speagle, *Mon. Not. R. Astron. Soc.* **493**, 3132 (2020).
  - [61] J. S. Speagle, *Mon. Not. R. Astron. Soc.* **493**, 3132 (2020).
  - [62] J. J. Blanco-Pillado, K. D. Olum, and B. Shlaer, *Phys. Rev. D* **92**, 063528 (2015).
  - [63] L. Lentati *et al.*, *Mon. Not. R. Astron. Soc.* **453**, 2577 (2015).
  - [64] S. A. Sanidas, R. A. Battye, and B. W. Stappers, *Phys. Rev. D* **85**, 122003 (2012).
  - [65] T. Charnock, A. Avgoustidis, E. J. Copeland, and A. Moss, *Phys. Rev. D* **93**, 123503 (2016).