

The International Pulsar Timing Array second data release: Search for an isotropic Gravitational Wave Background

J. Antoniadis,^{1,2,3} Z. Arzoumanian,⁴ S. Babak,^{5,6} M. Bailes,^{7,8} A.-S. Bak Nielsen,^{2,9} P. T. Baker,¹⁰ C. G. Bassa,¹¹ B. Bécsy,¹² A. Berthreau,^{13,14} M. Bonetti,^{15,16} A. Brazier,¹⁷ P. R. Brook,^{18,19} M. Burgay,²⁰ S. Burke-Spolaor,^{18,19,21} R. N. Caballero,²² J. A. Casey-Clyde,²³ A. Chalumeau,^{5,14,13} D. J. Champion,² M. Charisi,²⁴ S. Chatterjee,¹⁷ S. Chen,^{13,14} I. Cognard,^{13,14} J. M. Cordes,¹⁷ N. J. Cornish,¹² F. Crawford,²⁵ H. T. Cromartie,^{17,26} K. Crowter,²⁷ S. Dai,²⁸ M. E. DeCesar,²⁹ P. B. Demorest,³⁰ G. Desvignes,^{2,31} T. Dolch,^{32,33} B. Drachler,³⁴ M. Falxa,⁵ E. C. Ferrara,^{35,36,37} W. Fiore,^{18,19} E. Fonseca,^{38,19} J. R. Gair,³⁹ N. Garver-Daniels,^{18,19} B. Goncharov,^{40,41} D. C. Good,²⁷ E. Graikou,² L. Guillemot,^{13,14} Y. J. Guo,² J. S. Hazboun,⁴² G. Hobbs,⁴³ H. Hu,² K. Islo,⁴⁴ G. H. Janssen,^{11,45} R. J. Jennings,¹⁷ A. D. Johnson,⁴⁴ M. L. Jones,⁴⁴ A. R. Kaiser,^{18,19} D. L. Kaplan,⁴⁴ R. Karuppusamy,² M. J. Keith,⁴⁶ L. Z. Kelley,⁴⁷ M. Kerr,⁴⁸ J. S. Key,⁴² M. Kramer,^{2,46} M. T. Lam,^{34,49} W. G. Lamb,²⁴ T. J. W. Lazio,⁵⁰ K. J. Lee,^{22,51,2} L. Lentati,⁵² K. Liu,² J. Luo,⁵³ R. S. Lynch,⁵⁴ A. G. Lyne,⁴⁶ D. R. Madison,⁵⁵ R. A. Main,² R. N. Manchester,⁴³ A. McEwen,⁴⁴ J. W. McKee,⁵⁶ M. A. McLaughlin,^{18,19} M. B. Mickaliger,⁴⁶ C. M. F. Mingarelli,^{57,23} C. Ng,⁵⁸ D. J. Nice,⁵⁹ S. Osłowski,⁶⁰ A. Parthasarathy,² T. T. Pennucci,⁶¹ B. B. P. Perera,⁶² D. Perrodin,²⁰ A. Petiteau,⁵ N. S. Pol,²⁴ N. K. Porayko,² A. Possenti,^{20,63} S. M. Ransom,⁶⁴ P. S. Ray,⁶⁵ D. J. Reardon,^{7,8} C. J. Russell,⁶⁶ A. Samajdar,¹⁵ L. M. Sampson,⁴⁷ S. Sanidas,⁴⁶ J. M. Sarkissian,⁶⁷ K. Schmitz,⁶⁸ L. Schult,²⁴ A. Sesana,^{15,16} G. Shaifullah,^{15,16} R. M. Shannon,^{7,8} B. J. Shapiro-Albert,^{18,19} X. Siemens,^{69,44} J. Simon,^{50,70} T. L. Smith,⁷¹ L. Speri,³⁹ R. Spiewak,^{46,7,8} I. H. Stairs,²⁷ B. W. Stappers,⁴⁶ D. R. Stinebring,⁷² J. K. Swiggum,⁵⁹ S. R. Taylor,²⁴ G. Theureau,^{13,14,73} C. Tiburzi,¹¹ M. Vallisneri,^{50,74} E. van der Wateren,^{11,45} A. Vecchio,⁷⁵ J. P. W. Verbiest,^{9,2} S. J. Vigeland,⁴⁴ H. Wahl,^{18,19} J. B. Wang,⁷⁶ J. Wang,⁹ L. Wang,⁵¹ C. A. Witt,^{18,19} S. Zhang,⁷⁷ and X. J. Zhu⁷⁸

Affiliations are at the end of the paper

Accepted XXX. Received YYY; in original form ZZZ

ABSTRACT

We searched for an isotropic stochastic gravitational wave background in the second data release of the International Pulsar Timing Array, a global collaboration synthesizing decadal-length pulsar-timing campaigns in North America, Europe, and Australia. In our reference search for a power law strain spectrum of the form $h_c = A(f/1 \text{ yr}^{-1})^\alpha$, we found strong evidence for a spectrally-similar low-frequency stochastic process of amplitude $A = 3.8^{+6.3}_{-2.5} \times 10^{-15}$ and spectral index $\alpha = -0.5 \pm 0.5$, where the uncertainties represent 95% credible regions, using information from the auto- and cross-correlation terms between the pulsars in the array. For a spectral index of $\alpha = -2/3$, as expected from a population of inspiralling supermassive black hole binaries, the recovered amplitude is $A = 2.8^{+1.2}_{-0.8} \times 10^{-15}$. Nonetheless, no significant evidence of the Hellings-Downs correlations that would indicate a gravitational-wave origin was found. We also analyzed the constituent data from the individual pulsar timing arrays in a consistent way, and clearly demonstrate that the combined international data set is more sensitive. Furthermore, we demonstrate that this combined data set produces comparable constraints to recent single-array data sets which have more data than the constituent parts of the combination. Future international data releases will deliver increased sensitivity to gravitational wave radiation, and significantly increase the detection probability.

Key words: gravitational waves — pulsars:general — supermassive black holes — methods: data analysis — methods: statistical techniques

1 INTRODUCTION

Inspiralling supermassive black hole binaries (SMBHBs) with masses larger than $10^7 M_\odot$ are expected to generate the strongest gravitational-wave (GW) signals in the Universe. The incoherent superposition of all of these inspiralling SMBHBs should generate a stochastic GW background (GWB) that is the strongest in the nanohertz frequency band (e.g., Rajagopal & Romani 1995; Jaffe & Backer 2003; Sesana et al. 2008; Burke-Spolaor et al. 2019). Other sources that could also produce a stochastic background in the nanohertz band are cosmic strings (e.g., Ölmaz et al. 2010), cosmological phase transitions, and a primordial background produced by quantum fluctuations in the gravitational field in the early universe (e.g., Grishchuk 2005; Lasky et al. 2016). For comparison, the Laser Interferometer Gravitational-Wave Observatory (LIGO) and the Virgo Collaboration, which are terrestrial GW detectors and have detected GWs from merging stellar mass black holes and neutron stars (e.g., Abbott et al. 2019, 2021), are only sensitive to GW signals that are ten orders of magnitude higher in frequency than PTAs.

A nanohertz GWB can be detected using a precisely timed ensemble of millisecond pulsars (Sazhin 1978; Deweiler 1979), called a pulsar timing array (PTA, Foster & Backer 1990). The GWs distort the space-time between the Earth and pulsars, changing their proper distance, thereby leading to a measurable deviation of the pulsar pulse arrival times. Since such effects cannot be detected with confidence using only one pulsar, PTAs leverage the imprint of *spatially-correlated* timing deviations between pulsars which are separated by kiloparsec distances across the galaxy, yet are subject to the common influence of the GWB.

An isotropic GWB manifests itself as a long timescale, low-frequency (or red) common signal across the pulsars in a PTA. This common signal is characterized by the common spectrum and the inter-pulsar spatial correlations. For an isotropic GWB these spatial correlations are unique, referred to as the Hellings & Downs (1983) (HD) correlations, and thus are considered to be the “smoking gun” signature for the presence of a GWB (Tiburzi et al. 2016) in any PTA data set. The spectral amplitude of this common signal is determined by the characteristic strain, $h_c(f)$, of the GWB, which itself is a function of the physics sourcing the GWB (e.g., SMBHB masses, merger timescale, and number density) (e.g., Sesana 2013a; Kelley et al. 2017; Chen et al. 2019). Thus, precise spectral characterization of the GWB will allow us to extract the underlying astrophysics of the background, as well as distinguish between different sources of the GWB (e.g., Pol et al. 2021).

The ability to detect GWs relies on, among other things, the number of pulsars available to cross-correlate in GWB searches, and on the length of each pulsar data set (Siemens et al. 2013). Improvements in both of these parameters increases the detection significance of the GWB signal which in turn allows for better constraints on the parameters of the GWB spectrum (Pol et al. 2021). Hence, international efforts spanning decades from the European Pulsar Timing Array (EPTA, Desvignes et al. 2016), North American Nanohertz Observatory for Gravitational Waves (NANOGrav, Arzoumanian et al. 2016), and the Parkes Pulsar Timing array (PPTA, Manchester et al. 2013), as well as newer PTAs

such as the Indian Pulsar Timing Array (InPTA, Joshi et al. 2018), Chinese Pulsar Timing Array (CPTA, Lee 2016), and with the MeerKAT Interferometer in South Africa (Bailes et al. 2020) share and combine their data to form the International PTA (IPTA, Hobbs et al. 2010).

In this spirit of international collaboration the IPTA has produced two data sets to date. The first IPTA data release (DR1, Verbiest et al. 2016) consisted of 44 millisecond pulsars and yielded no conclusive detection of a GWB. The second IPTA data release (DR2, Perera et al. 2019) consists of 65 pulsars and is the focus of this analysis. The pulsars in DR2 have data sets spanning 0.5 – 30 years. For the first time, we process the data subsets from each individual PTA and search for a GWB in a self-consistent way, thus enabling us to make a fair comparison of respective PTA constraints.

Recently, a spatially uncorrelated (pulsar-weighted-average) spectrally similar common process or common-spectrum process (CP) was detected in the NANOGrav 12.5-year data set (Arzoumanian et al. 2020), the second data release of the Parkes Pulsar Timing Array (Goncharov et al. 2021a), and the EPTA six-pulsar data set of the second data release (Chen et al. 2021b). The process is modeled as an additional time-correlated term with the same power spectrum in all of the pulsars. However, there is little evidence to support the existence of spatial HD correlations in any of these data sets. We compare the IPTA DR2 constraints on the GWB with those obtained from these analyses.

The paper is organized as follows: in Section 2 we give an overview of the second IPTA data release, hereafter referred to as DR2. We describe our data analysis methods in Section 3, and give our results in Section 4. Caveats and implications of our analysis and results are discussed in Section 5, including the astrophysical interpretation of a potential GWB. The conclusion is given in Section 6.

2 IPTA DATA RELEASE 2

IPTA DR2 includes a combination of timing data from the following individual PTA data releases: the EPTA data release 1.0 (Desvignes et al. 2016), the NANOGrav 9-year data set (Arzoumanian et al. 2015), and the PPTA first data release (Manchester et al. 2013) and its extended version (Reardon et al. 2016). The EPTA data set includes high-precision timing data from 42 MSPs obtained with the largest radio telescopes in Europe – Effelsberg telescope, Lovell telescope, Nançay telescope, and Westerbork Synthesis telescope – covering data from 1996 to 2015 with a time baseline between 7–18 yr. In addition to these data, archival timing data of PSR J1939+2134 since 1994 was included. The NANOGrav 9-year data set includes high-precision timing observations from 37 MSPs obtained with the Robert C. Byrd Green Bank Telescope and the Arecibo telescope, spanning a time baseline between 0.6–9.2 yr, covering the data from 2004 to 2013. In addition, the long-term timing data of PSR J1713+0747 from Zhu et al. (2015) and the data of PSRs J1857+0943 and J1939+2134 from 1984 through 1992 (Kaspi et al. 1994) were included. The PPTA data set includes high-precision timing observations from 20 MSPs obtained with the Parkes radio telescope (also known as *Murriyang*) from 2004 to 2011. IPTA DR2 also included single frequency band (1.4 GHz/L-band) Parkes telescope

legacy data obtained since 1994. The additional 3.0 GHz timing data reported in Shannon et al. (2015) for PSRs J0437–4715, J1744–1134, J1713+0747, and J1909–3744 were also included in the data set. In total, the timing data from 65 MSPs were included in IPTA DR2, which has 21 more source than the IPTA DR1 (Verbiest et al. 2016). There are 27 and 7 MSPs in IPTA DR2 with a timing baseline >10 yr and >20 yr, respectively. All pulsars were observed at multiple frequencies. All EPTA and PPTA observations were averaged in time and frequency to obtain a single time-of-arrival (TOA) for each receiver and observation. The NANOGrav observations were averaged in time and included sub-band information, i.e., averaged in frequency to maintain a frequency resolution ranging from 1.5 to 12.5 MHz depending on the receiver and backend instrument combination, resulted in a single TOA for each frequency channel. More details about the constituent PTA data sets can be found in Perera et al. (2019).

The different data sets for a given pulsar in IPTA DR2 were combined by fitting for time offsets, referred to as JUMPs, in the timing model to account for any systematic delays between data sets. The highest weighted data set with the lowest sum of TOA uncertainties was used as the reference data set in this process. The timing models of pulsars included astrometric parameters, rotational frequency information, dispersion measure information, and Keplerian and Post-Keplerian parameters if the pulsar is in a binary system. For NANOGrav observed pulsars, “FD” parameters were included to minimize the effect of frequency-dependent profile variations of pulsars (see Arzoumanian et al. 2015). IPTA DR2 produced two data set versions depending on different methods of handling the dispersion measures (DM) variations of pulsars over time (VersionA and VersionB – see Perera et al. 2019, for details). In VersionA, the DM variations of pulsars were determined using DMModel described in Keith et al. (2013) and the noise parameters for different data sets were directly taken from their original data releases. In VersionB, the DM variations were modeled using the first two time derivatives of the DM and a time-correlated stochastic DM process in the timing model. The noise parameters were also re-estimated based on the new IPTA data combination in this version. We use VersionB for this work.

3 DATA-ANALYSIS METHODS

In this work we follow the conventions established by other pulsar timing array data analyses (i.e., Arzoumanian et al. 2016, 2020; Lentati et al. 2015; Goncharov et al. 2021a). The multivariate Gaussian likelihood $\mathcal{L}(\delta t|\theta)$ is employed to model noise and signal contributions, parametrised by θ , to the observed timing residuals. Our likelihood was of the same form as other PTA analyses, (e.g., Arzoumanian et al. 2015). We used `enterprise` (Ellis et al. 2020) to evaluate the likelihood and priors, and `PTMCMCSampler` (Ellis & van Haasteren 2017) to perform a Markov chain Monte Carlo (MCMC) simulation, drawing samples from the posterior probability distribution. Model selection was performed via the product-space sampling method (Carlin & Chib 1995; Hee et al. 2016). Additionally, we used the Savage-Dickey

approximation to the Bayesian evidence ratio when appropriate.

3.1 Noise models

For each pulsar, we modeled the TOAs with a combination of four processes: the timing model, white noise, intrinsic red noise, and DM variations. Deterministic contributions from the timing model, described in Section 2, were analytically marginalised (van Haasteren et al. 2009). The time-uncorrelated white noise was modeled with EFAC and EQUAD, and ECORR parameters for NANOGrav pulsars with their sub-banded TOAs (definitions of EFAC, EQUAD, and ECORR can be found in, e.g., Verbiest et al. 2016; Perera et al. 2019). Every observing receiver and backend system combination is given its own set of white noise parameters. The time-correlated red noise process (e.g., pulsar spin noise, Shannon & Cordes 2010) and stochastic DM variations (Keith et al. 2013) were modeled as Fourier basis Gaussian processes. In each case Fourier spectrum coefficients were modeled as power laws,

$$P_{\text{RN}}(f) = \frac{A_{\text{RN}}^2}{12\pi^2} f_{\text{yr}}^{-3} \left(\frac{f}{f_{\text{yr}}} \right)^{-\gamma_{\text{RN}}}, \quad (1)$$

$$P_{\text{DM}}(f, \nu) = \frac{A_{\text{DM}}^2}{12\pi^2} f_{\text{yr}}^{-3} \left(\frac{f}{f_{\text{yr}}} \right)^{-\gamma_{\text{DM}}} \left(\frac{1400 \text{ MHz}}{\nu} \right)^2, \quad (2)$$

where A is the power law amplitude, γ its spectral index, and $f_{\text{yr}} = 1 \text{ yr}^{-1} \approx 3.17 \times 10^{-8} \text{ Hz}$. The difference between these two processes lies in the radio frequency ν dependence. Intrinsic red noise is achromatic, i.e., frequency independent, while DM variations follow a ν^{-2} dependence (e.g., Lentati et al. 2016).

Despite all MSPs in IPTA DR2 exhibiting high rotational stability, such that the marginalized timing model, red, and DM noise terms are in general sufficient, certain pulsars have been found to experience timing events that need to be included in their data model. Of interest to this analysis is PSR J1713+0747, which was observed to experience multiple sudden drops in apparent DM with an exponential recovery (Demorest et al. 2013; Lam et al. 2018; Goncharov et al. 2021b). Only the first such event lies within the timespan of the IPTA DR2 and was included as an additional deterministic term to the full noise model of PSR J1713+0747. The amplitude, epoch and recovery time scale of the DM exponential dip are sampled simultaneously with the pulsar red and DM noise terms.

Lentati et al. (2016) found additional sources of red noise in IPTA DR1. These include radio frequency band-dependent and observing system-dependent terms, which may affect measurements of P_{RN} and P_{DM} , if not modeled. It is possible that mismodeling these effects can bias recovery of the CP. More prescriptive models for the CP should be less affected by this bias. Recent PTA analyses have included more complex red noise and DM variation models, where different pulsars in the array use different models (e.g., Aggarwal et al. 2019; Goncharov et al. 2021b). In the name of computational efficiency we opted to use the same power law models for all pulsars except when absolutely necessary, as is the case for PSR J1713+0747.

IPTA DR2, being the combination of data from multiple telescopes and many observing systems, has larger model

parameter space than its constituent data sets. The large number of model parameters and TOAs increases the computational complexity of the analysis. As we searched for long-term processes, such as the GWB, we limited our analysis to pulsars whose observation time exceeded 3 years. This reduced the number of pulsars from the full 65 in DR2 to 53. Additionally, we fixed the white noise parameters (EFAC, EQUAD, and ECORR) to median a posteriori values from single pulsar analyses. Both of these choices reduced the analysis parameter space to a more manageable size.

3.2 Common-spectrum process models

In addition to modeling noise intrinsic to the individual pulsars, we also include a red CP that is present in all of the pulsars. The source of this process could be the GWB, or any other common noise that manifests itself in all pulsars, such as clock errors (Caballero et al. 2016; Hobbs et al. 2020) or errors in the Roemer delays from Solar-system ephemeris (SSE) systematics (Tiburzi et al. 2016; Vallisneri et al. 2020). The choice of red noise priors also affects the recovery of a CP due to covariance between pulsar intrinsic red noise and the CP (Hazboun et al. 2020b; Goncharov et al. 2021a). Each of these effects can be distinguished by a unique pattern of spatial cross-correlations between pulsars. The cross-power spectral density is defined as,

$$S_{ab}(f) = \Gamma_{ab} P_{\text{CP}}(f), \quad (3)$$

where P_{CP} is the common-spectrum process and Γ_{ab} is the overlap reduction function (ORF) describing the inter-pulsar correlations.

For some analyses we did not account for any inter-pulsar correlations, taking $\Gamma_{ab} = \delta_{ab}$ to be the identity matrix. For others, we also included different choices of non-diagonal ORFs, such as the quadrupolar Hellings & Downs (1983) correlations that describe a GWB, dipolar correlations associated with SSE errors, or monopolar correlations, $\Gamma_{ab} = 1$, associated with clock errors. In some cases we split the diagonal auto-correlation part of the ORF from the off-diagonal cross-correlation part, treating them as independent processes as a consistency check. When modeling the CP using only the auto-correlations, it is possible to analyze the data from each pulsar independently, then recombine the results to achieve a joint posterior on the CP. We refer to this as the *factorized likelihood* approach.

We modeled the CP using a Fourier basis Gaussian process, using basis frequencies $f = 1/T, 2/T, \dots$, where T is the timespan between the earliest and latest observation in the data set. We model the power spectrum of the CP as a power law using Eqn. (1), replacing the pulsar noise amplitude and spectral index with those from the common process A_{CP} and γ_{CP} . In this parameterization of the power spectrum the characteristic strain spectrum for the GWB is

$$h_c(f) = A_{\text{CP}} \left(\frac{f}{f_{\text{yr}}} \right)^{(3-\gamma_{\text{CP}})/2}. \quad (4)$$

In some cases we fixed $\gamma_{\text{CP}} = 13/3$, equivalent to $\alpha = (3 - \gamma_{\text{CP}})/2 = -2/3$, the expected spectrum for a GWB composed of circular supermassive black holes (Phinney 2001), and in others we left γ_{CP} as a free parameter. To determine the number of Fourier frequencies used in the

power law CP model, we fit the power spectrum with a broken power law model. The broken power law is the sum of the standard, red power law and a white spectrum. This is implemented as a single spectrum with a fixed spectral index at low frequencies that smoothly transitions into a flat, white noise dominated spectrum at high frequencies:

$$P(f) = \frac{A_{\text{CP}}^2}{12\pi^2} f_{\text{yr}}^{-3} \left(\frac{f}{f_{\text{yr}}} \right)^{-\gamma_{\text{CP}}} \left[1 + \left(\frac{f}{f_{\text{bend}}} \right)^{1/\kappa} \right]^{\kappa \gamma_{\text{CP}}}, \quad (5)$$

where f_{bend} is the frequency where the spectral index of the power spectrum changes and κ controls the smoothness of the transition. In this model $P(f) \sim f^{-\gamma_{\text{CP}}}$ for $f \ll f_{\text{bend}}$, and $P(f)$ constant for $f \gg f_{\text{bend}}$. As a verification of our power law models, we performed a free spectral analysis, where the power at each frequency is fit independently rather than being constrained by a particular spectral shape, $P(f)$.

3.3 Frequentist analyses

As a comparison for our primary Bayesian data analysis pipeline, we performed a frequentist analysis using the noise-marginalized *optimal statistic*. The optimal statistic is an estimator for the amplitude of the GWB based on the inter-pulsar correlations (Anholm et al. 2009; Demorest et al. 2013; Chamberlin et al. 2015). Its original derivation assumed the pulsars have no intrinsic red noise. The noise-marginalized optimal statistic uses posterior samples from the Bayesian data analysis to marginalize over the pulsars' red noise. It has been shown to more accurately estimate the amplitude of the GWB when the pulsars have intrinsic red noise (Vigeland et al. 2018), as is the case in IPTA DR2.

4 RESULTS

The IPTA DR2 data set with its large number of pulsars (53 for this work), long timespan, and various independent observing systems offers a wealth of different analysis opportunities. Here we present a selection of analyses to give a complete picture of what IPTA DR2 teaches us and how it compares to other PTA data sets.

4.1 IPTA DR2 data set

We first show results from the full combined IPTA DR2 data set using methods which differ in their spectral modeling and choice of ORF. The full standard GWB search uses all the available information from the auto- as well as the cross-terms (which are assumed to follow the HD correlation). As the cross-terms, which come from the inter-pulsar correlations, are the defining feature of the GWB, we insist on their presence in order to confidently claim a detection. However, the auto-correlations are initially the dominant source of information, especially for spectral parameter estimation, and the detection of power in them is considered to be the first hint of a GWB (Romano et al. 2021). It is important to emphasize that detecting the auto-correlations alone is insufficient to claim a detection of a GWB.

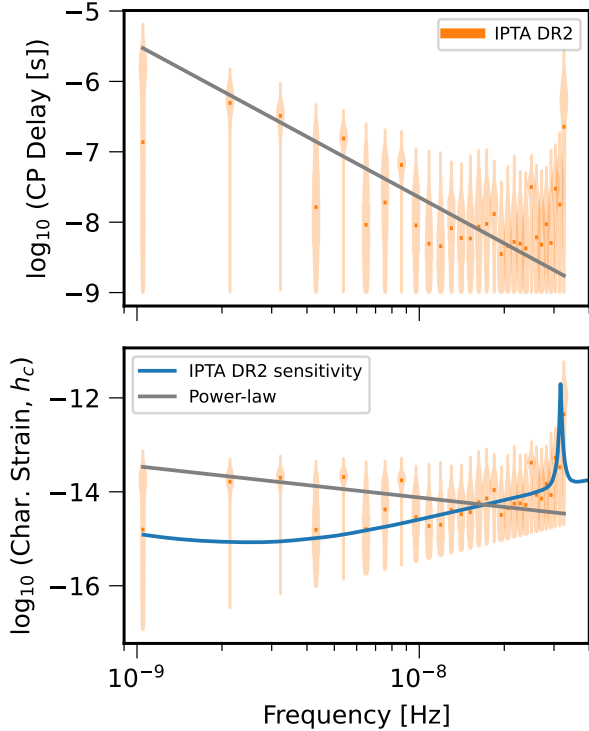


Figure 1. IPTA DR2 free spectrum and characteristic strain: The top panel shows the power in terms of time of arrival residuals at frequencies in the nanohertz band for the full PTA. The maximum likelihood power law is shown overlaid on posteriors for free spectral parameters, a generic model that measures power at various frequencies without imposing any empirical model. The bottom panel shows the power law and free spectral information from above converted into units of characteristic strain, i.e., the noise power measured in the same units as GW amplitude. The additional line shows the characteristic strain for the detector using the noise parameters for the pulsars. The lower limit of all violins is a result of the lower bound of the prior range for each frequency component.

4.1.1 Common-spectrum process

To begin, we apply a free spectral model for the CP, measuring the amount of power at each sampling frequency independently, up to 30 frequency bins, the result of which is shown in the top panel of Figure 1. The common red noise power can be converted into GW strain using Eqn. (4). This alternative representation of the IPTA DR2 analysis can be found in the lower panel of Figure 1. For reference, we include the predicted sensitivity curve, made using *hasasia* (Hazboun et al. 2019a,b) and the measured white noise parameters of the DR2 data set. Note that the noise power spectral density used in this curve only contains TOA errors, EFAC, EQUAD and ECORR, and does not contain any estimates for the red noise, as for many pulsars it is difficult, and in fact the point of this analysis, to disentangle intrinsic red noise from a GWB. Hence the low-frequency end of the sensitivity represents a “best case” scenario for comparison. The lowest frequency bin corresponds to the longest timespan, and only two pulsars, J1939+2134 (29 yr) and J1857+0943 (28 yr), have observation baselines suffi-

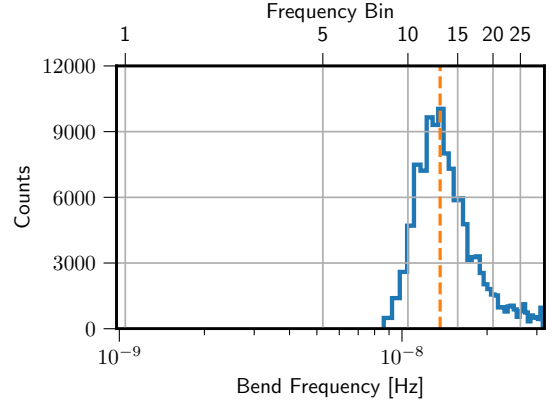


Figure 2. Bend Frequency: The posterior for the bend frequency parameter in a broken power law search is shown. The peak of the posterior is at the 13th frequency, $13/T$ for the data set, denoted by the dashed vertical line.

cient to probe this frequency. However, both have significant RN with spectral indices ~ 3.3 (Perera et al. 2019). Therefore, it is not surprising that we do not confidently detect any power there, evidenced by the wide tail extending to low power and median of $\sim 10^{-7}$ s. However, the second, third, fifth, and eighth frequency bins show power well above the expected sensitivity of IPTA DR2. This could either be the emergence of a GWB or some other unmodeled noise process.

The CP power spectrum can be modeled with a simple power law using Eqn. (1). Arzoumanian et al. (2020) have shown that the choice of the number of modeled frequencies can affect the constraints on the power law amplitude and spectral index. Thus, we apply the broken power law model from Eqn. (5) to find the optimal number of frequency bins for the analysis. Figure 2 shows the marginalized posterior on the bend frequency. We can identify a clear peak at the 13th frequency in N/T , corresponding to a frequency of 1.4×10^{-8} Hz, indicated by the orange dashed line. For the remainder of this work, we will limit the search to use only the lowest 13 frequencies with the simple power law model. This produces constraints equivalent to an analysis using the broken power law, but in a simpler and computationally efficient way.

We have also verified that the addition of the *BAYESEPHM* SSE model (Vallisneri et al. 2020) to the analysis does not change our results significantly from an analysis that fixed the SSE to DE438. The nearly 30 years of timespan allows for the separation of the SSE effects from other correlated signals (Vallisneri et al. 2020). For simplicity, we will only show results with DE438, unless stated otherwise.

Figure 3 compares the results when using two different ORFs. The model that uses only the auto-correlation terms, which we denote CP in Table 1, is very strongly favored over a model with only intrinsic pulsar noise and no common-spectrum process with \log_{10} Bayes factor of 8.2. Despite the large Bayes factor in favor of the CP, this does not suffice to claim a GWB detection, as we have only used the auto-correlations. This strong evidence only indicates that a number of pulsars have red noise with similar spectral characteristics. We must turn to the cross-correlations

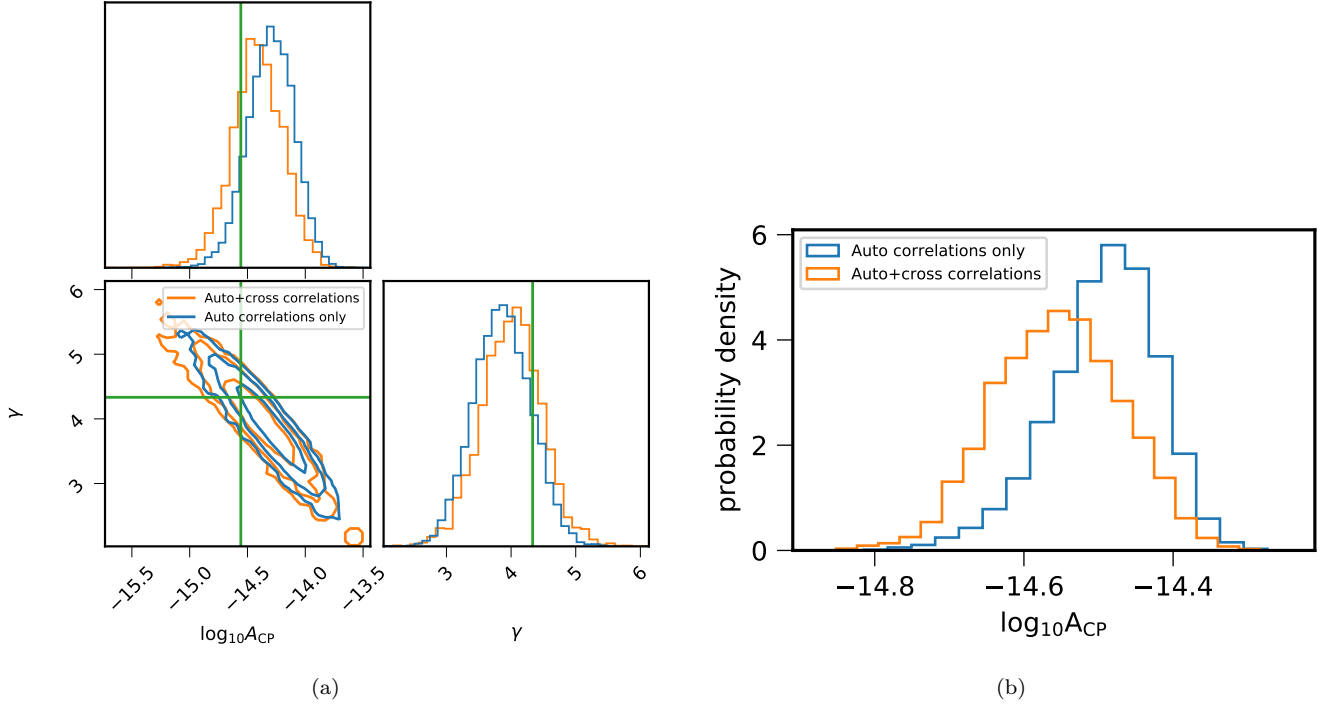


Figure 3. Comparison of common-spectrum process parameters when using auto-correlations only and the full auto+cross-correlated HD model. *left*: 2D posterior for common-spectrum process power law parameters. Green lines mark $\gamma = 13/3$ and $A_{CP} = 2.8 \times 10^{-15}$, while the contours represent the 1-, 2- and 3- σ confidence intervals. *right*: 1D posterior for common-spectrum process power law amplitude, using fixed spectral index $\gamma = 13/3$.

Table 1. Bayes factors model comparison: The table shows the logarithmic Bayes factors for a number of model comparisons from the hypermodel and factorized likelihood (marked with an asterisk*) analyses. The preferred model is on the left side of the two models. Brackets indicate the uncertainty in the last digit of the Bayes factors.

Model comparison	\log_{10} BF
HD vs CP	0.3111(6)
CP vs Pulsar Noise	8.2*
CP vs Monopole	4.67(2)
CP vs Dipole	2.28(3)

to determine if this CP is HD correlated as a GWB should be. Using the full HD ORF containing both auto- and cross-correlations, we find only middling evidence in favor of the auto+cross HD model. The \log_{10} Bayes factor for the full HD model compared to the auto-correlated only CP is 0.3, as shown in Table 1.

Figure 3a shows the 2D posterior contours of these two models are in relatively good agreement. A small shift towards lower amplitudes and higher spectral index can be seen when using the full HD ORF with both auto- and cross-correlations. Using the auto-correlation terms only, we find $A_{CP} = 5.1^{+6.7}_{-3.1} \times 10^{-15}$ and $\gamma_{CP} = 3.9 \pm 0.9$, where the errors represent 95% credible regions. Using the full HD ORF we can constrain the CP power law to $A_{CP} = 3.8^{+6.3}_{-2.5} \times 10^{-15}$ and $\gamma_{CP} = 4.0 \pm 0.9$.

When we fix the power spectrum index to $\gamma = 13/3$ as shown in Figure 3b, it is clear that full HD model finds a systematically lower amplitude. In this case we find an am-

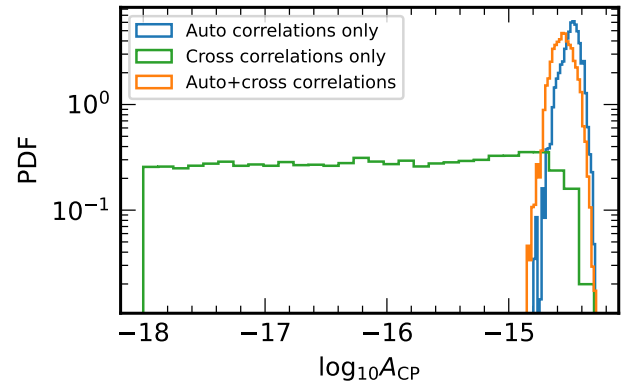


Figure 4. The constraints from the split ORF analysis on the common process amplitude with spectral index fixed to $\gamma = 13/3$. The posterior from the auto-correlations is well constrained, while the posterior from the cross-correlations is unconstrained, but prefers amplitudes slightly smaller than the auto-only analysis. The effect of this is seen in the amplitude posterior from the full auto+cross-correlation model, where the posterior peaks at lower amplitude than the auto-only analysis.

plitude of $A_{CP} = 3.2 \pm 1.0 \times 10^{-15}$ for the auto-correlation only analysis and an amplitude of $A_{CP} = 2.8^{+1.2}_{-0.9} \times 10^{-15}$ using the full HD ORF, where the uncertainties represent the 95% credible regions. These results are in broad agreement with published constraints on the CP. A more detailed comparison can be found in Section 4.3.

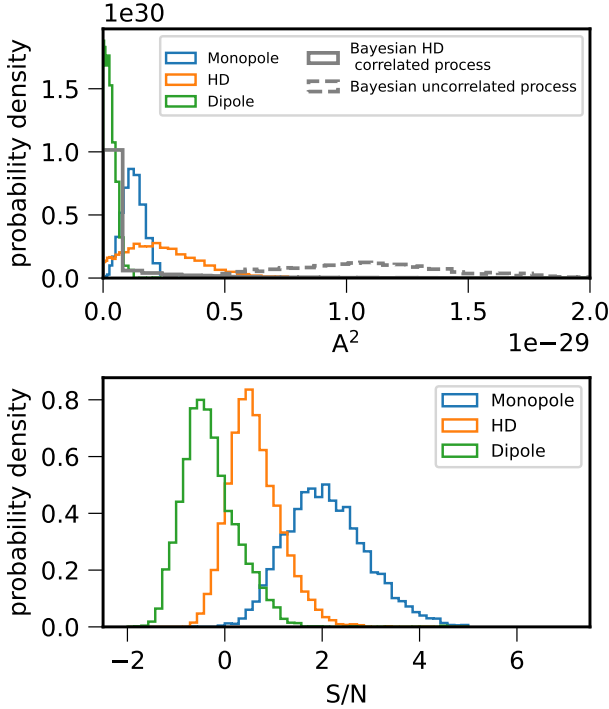


Figure 5. Results from the noise marginalized optimal statistic. *Top:* Optimal statistic, A^2 , distribution for monopole, dipole and HD ORFs. The relevant posteriors from the Bayesian split ORF analysis are also shown for comparison. *Bottom:* The signal-to-noise (S/N) distribution for monopole, dipole and HD ORFs.

4.1.2 Split ORF analysis

Similar to how we may consider the auto-correlation parts of the ORF alone, the full ORF can be split into two independent processes. In this case the auto-correlation and the cross-correlation parts each have their own independent amplitude, as was done in [Arzoumanian et al. \(2020\)](#). In the HD ORF the auto-correlation part is $\Gamma_{aa} = 1$ and the cross-correlation parts are suppressed by at least a factor of 2, $\Gamma_{ab} < 0.5$. This makes the cross-correlations harder to constrain. [Figure 4](#) shows the posteriors for the two amplitudes of a split ORF analysis for fixed $\gamma = 13/3$, compared to the full auto+cross-correlation model. The cross-correlations do not have sufficient precision to place constraints on the amplitude of the GWB. However, they do place a 95% upper limit of 3.6×10^{-15} on the GWB when the prior choice is taken into account. This is consistent with the amplitude derived using the full auto- and cross-correlation model in [Fig. 4](#). The auto-correlation terms are much more informative. Combining the information from both shifts the amplitude towards lower values. This shows that the cross-terms can contribute to the full GWB search, even if they provide less information. The auto-correlations are more likely to be affected by intrinsic pulsar noise. Using a more sophisticated noise model for each pulsar can help produce a more robust estimate on the amplitude of any CP.

Table 2. The p -values calculated from various false alarm analyses of the data set. The measured values of the S/N are compared to the distribution of 10k analyses where the correlations are broken with phase shifts and sky scrambles. Since a monopolar spatial correlation is uniform across the sky, sky scrambles are unable to break the correlations. Hence only the phase shift p -value is quoted.

	p , phase shift	p , sky scramble
HD ORF	0.25	0.24
Monopole	0.03	—

4.1.3 Optimal statistic

[Figure 5](#) shows the amplitudes and signal-to-noise ratio (S/N) that are recovered by the pulsar noise marginalized optimal statistic (OS) method, which uses cross-correlations only. We find no evidence for a dipolar correlated process, as the amplitude and S/N for this model are centered on 0. SSE systematics are expected to manifest at specific frequencies related to the celestial bodies. The IPTA DR2 data set is long enough to probe lower frequencies that should be less affected by SSE errors ([Vallisneri et al. 2020](#)). The $S/N = 0.6^{+1.2}_{-0.8}$ for the Hellings-Downs correlation is insufficient to claim a detection. This is consistent with the Bayesian model selection. The HD amplitude from the OS seems to be in tension with the Bayesian results for the auto-correlated CP, but consistent with the Bayesian results for the full HD model. This strengthens the case that the cross-terms have a significant role to play in parameter estimation as well as detection confidence. Finally, the OS has the largest $S/N = 2.0^{+1.8}_{-1.4}$ for a monopole with a small amplitude. This can be due to the complexity of IPTA DR2 and some amount of unmodeled noise.

As the spatial correlations are not well constrained, see [Figure 6](#), both the HD and monopolar correlation can fit the data. We have binned pulsar pairs according to their angular separation. Increasing the number of pulsars in the array as well as better timing of pulsars can help to tighten the constraints.

We test the significance of the OS S/N by performing two analyses which estimate the false alarm rate for a given S/N. So called phase-shifts and sky scrambles ([Cornish & Sampson 2016; Taylor et al. 2017](#)) break the correlations between the pulsars leaving the red noise power in the pulsars, but removing evidence for spatial correlations. By analyzing the phase-shifted and sky scrambled data we can determine the rate of observing a particular S/N for a type of correlations in data that has none. These two false alarm studies result in p -values, [Table 2](#), too high to conclude that there is evidence for any correlations. It is possible that the measured HD S/N can therefore arise by chance from a common process with no spatial correlation.

4.2 IPTA DR2 data subsets

With the large volume of the IPTA DR2 data set we can look at different subsets to investigate hints of the origin and evolution of the CP signal across different slicings of the full data set.

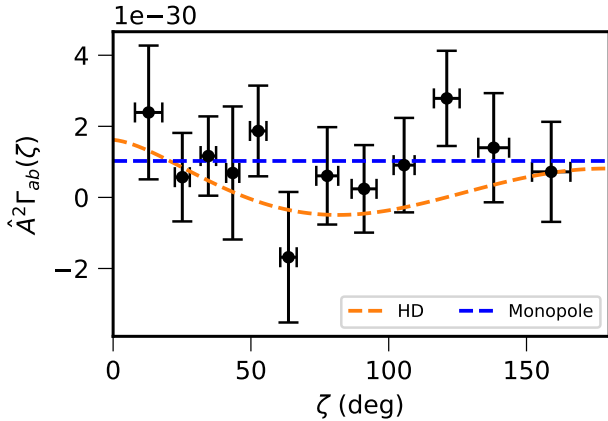


Figure 6. Cross-correlation ORF curve from the optimal statistic. The black points indicate the amount of cross-correlation for a given angular separation. Due to the large number of pulsar pairs, we have binned multiple pairs with similar angular separation. The blue and orange dashed lines show the best-fit values for the HD and Monopole correlations.

4.2.1 Pulsar-based selection

Since a PTA is made from a number of single pulsars, we can look at how each pulsar contributes to the CP by itself. The dropout factor gives a measure on how consistent a given pulsar’s intrinsic red noise is with the CP by comparing a model with and without the CP for the pulsar (see e.g., [Arzoumanian et al. 2020](#)). The dropout factors for each pulsar computed using both the traditional hypermodel and factorized likelihood approaches are shown in [Figure 7](#). About 20 pulsars have factors > 1 , while only three slightly disfavor the CP, with the remaining pulsars displaying indifference.

Monte Carlo sampling uncertainties on the dropout factors (computed either way) can be estimated through statistical bootstrapping ([Efron & Tibshirani 1994](#)). In the hypermodel dropout analysis the MCMC chain is re-sampled with replacement, generating a new statistical realization of the sampled chain that is exceptionally unlikely to be identical to the original chain. This process was repeated 10^3 times, generating many realizations of the MCMC chain from which the dropout factors were computed. Hence a distribution of dropout factors over bootstrap realizations was generated for each pulsar, allowing us to compute median values and 95% confidence intervals. A similar procedure was performed for the factorized likelihood approach. For a given bootstrap realization, each individual pulsar’s MCMC chain was re-sampled with replacement. With the re-sampled CP posteriors for each pulsar, the factorized-likelihood approach pieces together the dropout factor by iteratively removing pulsars from the array, and all by using bootstrapped pulsar CP posterior chains. This process was repeated for 10^3 bootstrap realizations across 25 different combinations of meta-parameters used in the factorized-likelihood dropout factor calculation. The end result is that the median dropout factor and 95% confidence intervals were computed from a total of 2.5×10^4 factorized-likelihood dropout values for each pulsar. As is seen in [Figure 7](#), all dropout factors are consistent between both techniques. The vast majority of pulsars have dropout factors with overlapping error bars from both

methods, and those that don’t are within a few sigma of each other. Those pulsars that show the largest disparity are ones for which there were MCMC sampling inefficiencies that manifested in different stages of the dropout factor calculation, e.g., in the Savage-Dickey density ratio, or in the integral of the $(N - 1)$ array’s CP likelihood over the posterior of a given pulsar (Taylor et al., in prep.).

The modularity and speed allowed by the factorized likelihood method can be used to approximate different combinations of pulsars within the array. These sub-arrays are a useful way to verify and understand the results we see in the full array. We created four sub-arrays, consisting of pulsars with the highest/lowest dropout factors, longest/shortest timespans. These pulsars were selected by sorting all pulsars in the array by their dropout and time-span characteristics, and then taking the top half of 27 pulsars and the bottom half of 26 pulsars. The Savage-Dickey density ratio was calculated for these sub-arrays to compare them to that of the full array. The sub-array made up of the top half of pulsars according to dropout factor had a Savage-Dickey density ratio of 5.6×10^9 , an order of magnitude larger than that of the unaltered array, 1.6×10^8 . The corresponding sub-array of the bottom half of pulsars based on dropout had a density ratio of 1.8. When comparing the pulsars with the longest and shortest timespans, the sub-array with the longer timespans had a Savage-Dickey density ratio of 1.8×10^7 and the shorter timespan sub-array had a density ratio of 3.6. These results were not surprising, as the dropout factor is a method of measuring the evidence for the array’s common process in a particular pulsar, so by removing those that have low (high) dropout factors, the evidence for the CP will increase (decrease). Similarly, pulsars with short timespans are not sensitive to the lowest frequencies explored by the array when in combination with longer timespan pulsars.

4.2.2 Splitting IPTA DR2 by time

To test the evolution of the common-spectrum process, the DR2 is split into two data sets that have equally long timespans, i.e., cutting the DR2 in two time slices (in a similar manner as [Hazboun et al. 2020a](#)). The two data sets are not fully equivalent though: the early part contains only 19 pulsars and is mostly dominated by single-radio frequency observations, while the second part has data from all 53 pulsars as well as multi-radio frequency coverage and higher quality timing measurements. Each data set is then analysed separately. We find that the first half gives little information to the CP, with a broad power law 2D posterior contour that still encompasses the contour from the full data set. The second half contains the majority of information and produces almost identical constraints as the full data set. This is the expected evolution as the quality of the data set gradually improves over time ([Hazboun et al. 2020a](#)).

4.2.3 Constituent data sets

We can also select the data that were provided by the constituent PTA collaborations to get three data subsets: EPTA, NANOGrav and PPTA. As each data subset has a different timespan, we set a frequency cutoff at 1.4×10^{-8} Hz to limit the number of frequencies for the analyses. [Figure 8](#)

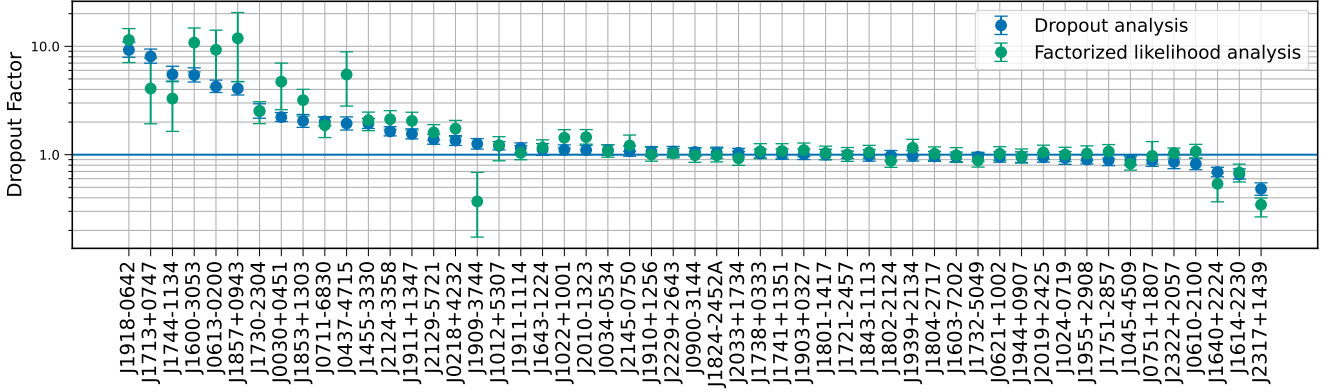


Figure 7. Individual pulsar consistency with common-spectrum process, error bars represent 95% credible intervals. Pulsars with dropout factors > 1 contribute to the detection of the CP. Dropout factors of ~ 1 correspond to no evidence for or against the CP, usually due to higher white noise levels and/or shorter observation timespans. Pulsars with dropout factors < 1 are in tension with the CP.

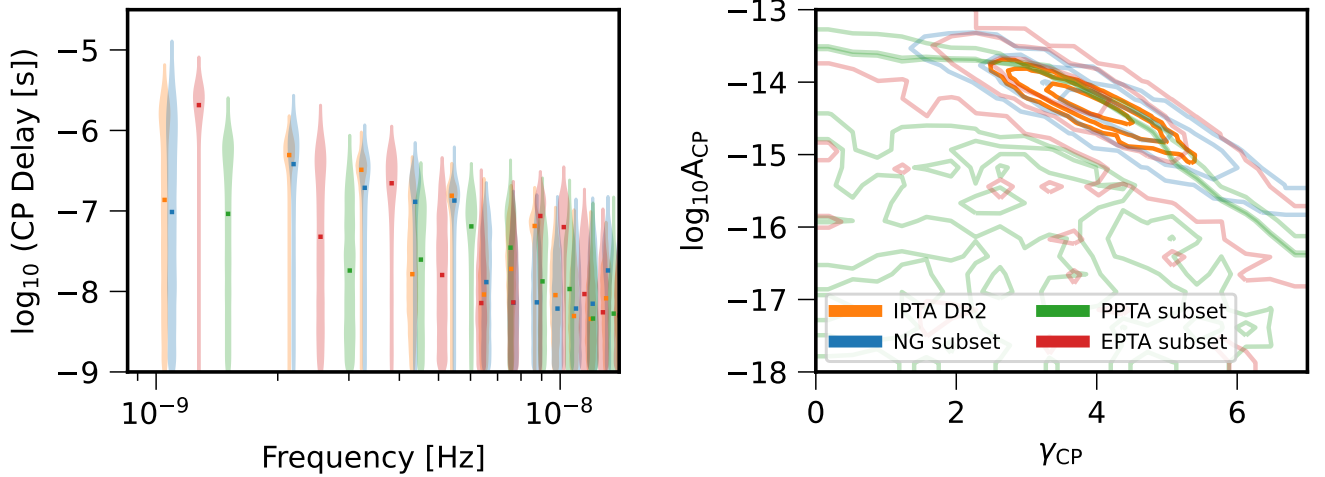


Figure 8. Comparison of IPTA DR2 to constituent data sets. *left:* Free spectral common-spectrum process model. The different sampling frequencies are a result of the constituent subsets covering different timespans. In all cases the lowest frequency is the inverse observation time. *right:* 2D posterior for CP parameters log-amplitude and spectral index, where the contours represent 1-, 2-, and 3- σ confidence intervals. The combined IPTA DR2 data set is more constraining than its parts.

shows that IPTA DR2 produces the tightest constraints on the CP power law compared to the constituent data sets. While the PPTA data is still consistent with an upper limit, some support for a common red noise can be found with the EPTA and NANOGrav data. The free spectra also show consistency with a power law model that spans across all three constituent data sets and IPTA DR2.

4.3 Comparison with other recent data sets

Since the data from the regional PTAs were combined to form the IPTA DR2, the regional PTAs have continued to collect data and improve their data analysis methodology. We can compare the results using the older IPTA DR2 data set and the most recent data sets from NANOGrav, PPTA and EPTA. Compared to the constituent PTA data sets, the recent NANOGrav data set includes ~ 4 more years and 10 new pulsars, the PPTA expands by ~ 3 years and 7 pulsars,

Table 3. Mahalanobis distance between CP parameters (log-amplitude and spectral index) for each pair of PTAs. For all cases, there is less than 3-sigma separation.

	EPTA	PPTA	NANOGrav
IPTA	0.6	2.6	2.6
EPTA	–	2.3	2.4
PPTA	–	–	1.4

the EPTA DR2 adds ~ 7 years for 6 pulsars (Alam et al. 2021; Kerr et al. 2020; Chen et al. 2021b).

The published free spectral and power law model recoveries can be found in Figure 9. For simplicity, we also show the recovered amplitudes at the reference frequency of $1/(1\text{yr})$ and fixed $\gamma_{\text{CP}} = 13/3$ in Figure 10. The Mahalanobis distance D_M acts as a generalization to compute the n-dimensional sigma deviation between two distributions

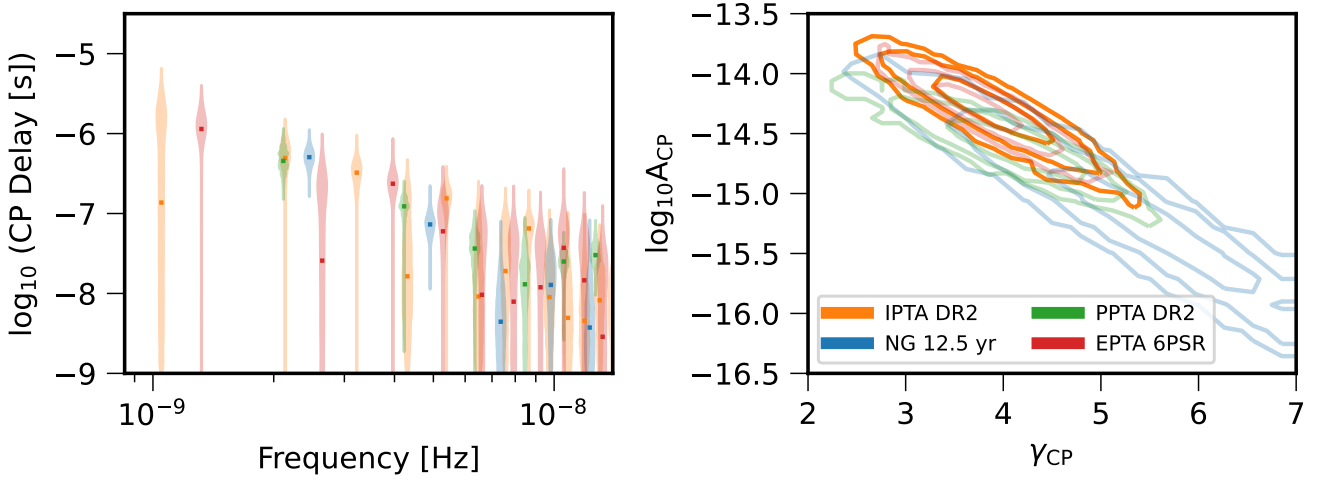


Figure 9. Comparison of IPTA DR2 to other recent data sets. *left*: Free spectral common-spectrum process model. The inclusion of legacy data not used in recent PTA analyses allows IPTA DR2 to reach lower frequencies despite missing the most recently collected data. *right*: 2D posterior for CP parameters log-amplitude and spectral index, where the contours represent the 1-, 2-, and 3- σ confidence intervals. All recent data sets are in broad agreement on the characteristics of a common-spectrum process.

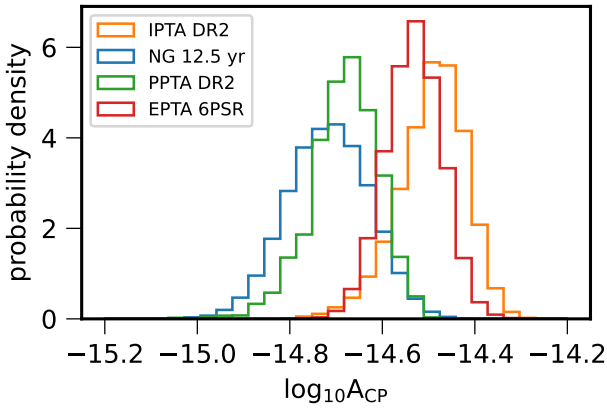


Figure 10. CP amplitude posteriors for fixed spectral index, $\gamma = 13/3$. IPTA DR2 and EPTA DR2 find a systematically higher amplitude for the common-spectrum process than NANOGrav 12.5 yr and PPTA DR2, although the disagreement is not substantial.

(Mahalanobis 1936),

$$D_M = \sqrt{(\vec{\mu}_1 - \vec{\mu}_2)\Sigma^{-1}(\vec{\mu}_1 - \vec{\mu}_2)}, \quad (6)$$

where $\vec{\mu}_1$ and $\vec{\mu}_2$ are the mean vectors of the multivariate distributions to be compared and $\Sigma = \Sigma_1 + \Sigma_2$ is the joint covariance. To quantify the overlap and consistency of the power law parameters as determined using each dataset, the Mahalanobis distance between the 2D posterior distributions are computed in Table 3. Despite some differences the posteriors overlap better than 3-sigma for all pairs of distributions.

IPTA DR2, using older observations, still shows similar features as the NANOGrav 12.5, 6-pulsar EPTA DR2 and PPTA DR2 analyses, which have added a significant amount of new data to the regional PTA data sets. A future combination of these data sets will boost the total PTA sen-

sitivity in the same way IPTA DR2 is more sensitive than its constituent data sets. Future combined IPTA data sets will be important for investigating the origin of this common-spectrum process.

5 DISCUSSION & OUTLOOK

5.1 Source of the common-spectrum process

The first IPTA data release did not show signs of a common-spectrum temporally-correlated process, but set an upper limit of 1.7×10^{-15} instead. This appears to be in tension with our results from analysis of the second data release with a CP amplitude of 2.8×10^{-15} . However, there are two major differences to point out: 1) the different choice of priors for the pulsar red, DM and common noise (Hazboun et al. 2020b) and 2) the DR1 upper limit was computed without the use of a SSE uncertainty model (Vallisneri et al. 2020). Both of which have been shown to lead to an increase in the upper limit, alleviating tensions between the DR1 and DR2 CP amplitudes.

As in other recent PTA analyses, we find strong evidence in favor of the CP over the noise only hypothesis. It is important to note that 1) the lack of support for GW-like spatial correlations prohibits any claims of GW detection, however 2) this type of evidence for a similar red noise is expected to precede a detection of spatial correlations (Siemens et al. 2013; Pol et al. 2021; Romano et al. 2021).

Goncharov et al. (2021a) recently demonstrated that the common-spectrum process model is favored over the noise-only hypothesis when the noise spectra cluster in a similar range, and it is not favored anymore when the noise spectra are drawn from the prior distribution. Because we know that the employed prior distribution for red noise parameters is not representative, it is possible that the evidence we find for a common-spectrum process is caused by a rejection of a null hypothesis rather than by all pulsars exhibiting the spatially-uncorrelated component of a GWB.

Thus, it is important to examine the single pulsar red noise in detail. We have looked at constraints on the simple power law models for the pulsars used in the CP search. In general, pulsars with detectable intrinsic noise have comparable or larger noise than the CP; pulsars without red noise typically have large amount of white noise, such that the CP is ‘hidden’. One noticeable exception is PSR J2317+1439, whose noise spectrum falls clearly below the CP, see also its low dropout factor in Figure 7.

As the search for the common spectrum can be influenced by pulsar intrinsic noise, especially in an inhomogeneous data set, the crucial analysis has to consider information from the cross-correlations. It should be noted that the median amplitudes are slightly different in the analyses with and without spatial correlations, 2.8×10^{-15} vs 3.2×10^{-15} . One can also note the stark difference in the posterior for the split ORF analysis, Figure 4 and the optimal statistic analyses vs the Bayesian uncorrelated analysis, Figure 5. In other analyses (Arzoumanian et al. 2020; Goncharov et al. 2021a; Chen et al. 2021b) the amplitudes between the two analyses are more in line with one another. The difference here could be in part due to the very long baselines of only a handful of pulsars. This legacy data allows only scant opportunity for correlations amongst those few pulsars, while the long baselines allow the detection of auto-correlated power, even in noisier data. Another possibility is that there is unaccounted for noise in individual pulsars that is contaminating the signal. Advanced models to take these pulsar noises into account for the GWB search have been shown to affect the individual pulsar red noise (e.g., Arzoumanian et al. 2020; Goncharov et al. 2021a, Chalmereau et al., in prep.).

5.2 Other Correlated Signals

Spatial correlations in pulsar timing data have been studied in depth in the literature (Tiburzi et al. 2016; Roebber 2019), and their consideration is an important part of any GW detection procedure. While GWs induce a quadrupole-dominated set of correlations there are other types of spatial correlations between pulsar data sets (Roebber & Holder 2017; Tiburzi et al. 2016; Roebber 2019). Monopolar spatial correlations, i.e., all pulsars seeing the same shifts in residuals irrespective of sky position, can manifest from clock errors, either in the BIPM clock standards, or the various observatory clocks used across the world (Hobbs et al. 2020). Dipolar spatial correlations can manifest from the error in measurement of processes where the motion of the Earth in the solar system is important (Caballero et al. 2018). This is most direct in the modeling of the solar system barycenter frame of reference, into which all pulsar TOAs are transformed. Errors in solar wind modeling can also add dipolar correlations (Tiburzi et al. 2016). While monopole and dipole correlations are theoretically orthogonal to HD correlations in the space of overlap reduction functions, real data with noise can result in some of these modes mixing (Roebber 2019). This mixing could erroneously be detected as a GWB.

The polarization content of the GWB could also deviate from the two tensor transverse (TT) modes predicted by general relativity which lead to the HD spatial correlations (Arzoumanian et al. 2021b). Deviations from general relativity would result in a correlation pattern that differs

from HD. We would like to emphasize that the current data set does not allow us to draw any conclusions on the presence of spatial correlations. As can be seen in Figure 6 the uncertainties on the spatial correlation coefficients $\Gamma_{ab}(\xi)$ determined by the optimal statistic analysis are large. For most angular bins the correlation is indistinguishable from zero, corresponding to the uncorrelated CP. Close to submission of this paper, we noticed a preprint by Chen et al. (2021a). They have analyzed the IPTA DR2 searching for alternative GW polarizations and claim evidence for spatial correlations induced by a scalar transverse (ST) polarization mode. Chen et al. (2021a) also report a Bayes factor in favor of HD correlations (TT modes) compared to an uncorrelated CP about six times larger than we find: 12 against our 2. Even though we have not searched for the ST mode, we would like to highlight that the reported high evidence of spatial correlations in Chen et al. (2021a) is contrary to what we conclude using the same data set. The scalar transverse ORF is positive definite and should be accompanied by positive evidence for a monopolar correlation in analyses using only the cross-terms (Arzoumanian et al. 2021b). This is the case in our optimal statistic analysis where the monopolar correlations have the largest S/N and smallest false alarm p -value. In this sense finding some evidence in favor of ST correlations is not too surprising, however, we find no conclusive evidence in favor of any correlation pattern. Using more information in our analysis with both auto- and cross-terms disfavors monopolar correlations compared to an uncorrelated CP. Therefore, the conclusions of Chen et al. (2021a) need to be taken with caution. Finally, Chen et al. (2021a) find a Bayes factor in favor of the common process to be several orders of magnitude smaller ($\log_{10} \text{BF} = 4.6$ compared to 8.2) than we do. They use a different pulsar noise model, including an additional sinusoidal annual DM variation in all pulsars, which could account for some, but likely not all, of the differences.

5.3 Astrophysical Interpretation

The nHz GWB is generally thought to be dominated by GW emission from SMBHBs (Burke-Spolaor et al. 2019), with the most massive local SMBHBs expected to be individually observable in the next 5-10 years (Mingarelli et al. 2017). Given that these systems are just a local subset of the cosmological SMBHB population producing the GWB, their local number density, $\Phi_{\text{BHB},0}$, should correlate with the amplitude of the GWB. The GWB amplitude induced by a cosmological population of circularized SMBHBs can be expressed in geometric units (where $G = c = 1$) as (e.g. Phinney 2001; Sesana et al. 2008; Sesana 2013b)

$$A_{\text{GWB}}^2 = \frac{4}{3\pi^{1/3}} \iiint dM_1 dz dq \frac{\mathcal{M}^{5/3}}{(1+z)^{1/3}} \frac{d^3\Phi_{\text{BHB}}}{dM_1 dz dq}, \quad (7)$$

where M_1 is the mass of the primary SMBH, M_2 the mass of the secondary, $q = M_2/M_1 \leq 1$ is the mass ratio, $\mathcal{M}^{5/3} = [q(1+q)^{-2}]^{5/3} M^{5/3}$ is the SMBHB chirp mass with total binary mass $M = M_1 + M_2$, and $d^3\Phi_{\text{BHB}}/(dM_1 dz dq)$ is the differential comoving number density of SMBHBs per unit M_1 , z , and q .

To determine the local number density of SMBHBs implied by a $A_{\text{GWB}} \approx 2.8 \times 10^{-15}$ background we use the quasar-based SMBHB model from Casey-Clyde et al. (2021),

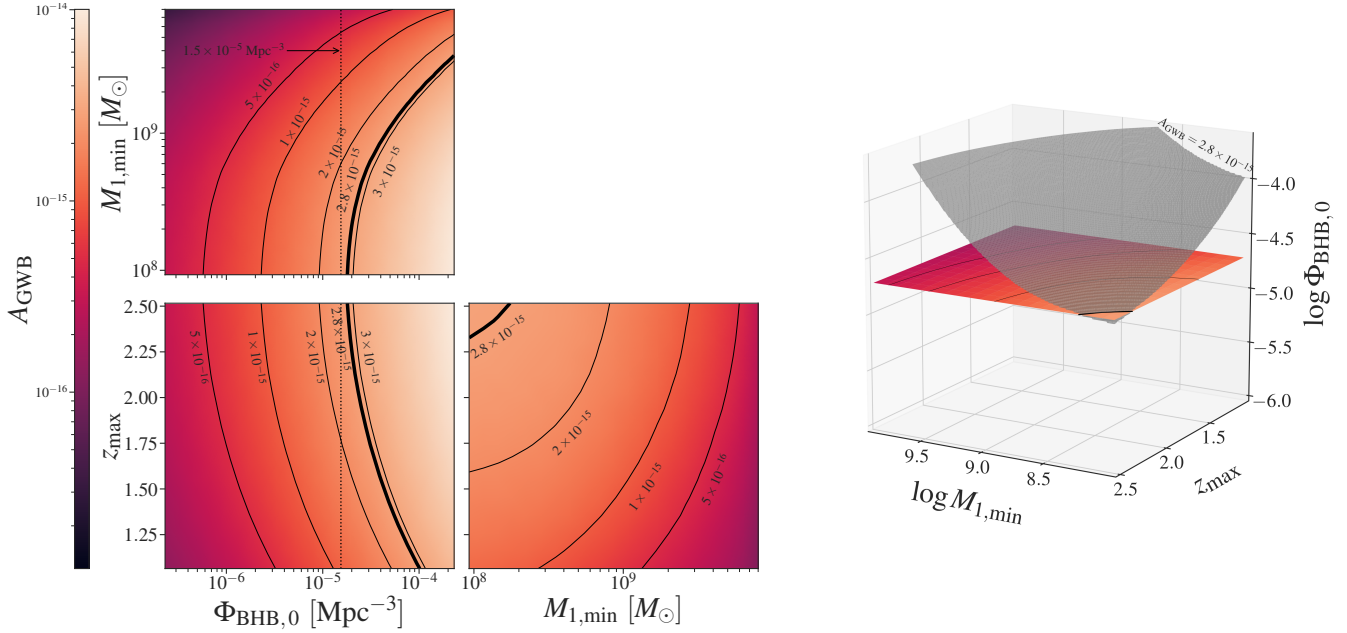


Figure 11. The GWB characteristic strain as a function of the local SMBHB number density, $\Phi_{\text{BHB},0}$, and the minimum primary BH mass, $M_{\text{BH},1,\text{min}}$, and maximum redshift, z_{max} , of the population contributing $\gtrsim 95\%$ of the GWB signal. *Left:* Three representative slices of the strain in this parameter space (one along each axis), with solid contours showing their intersection with isosurfaces of constant strain value (A_{CP} shown in bold). *Right:* 3D visualization of the $z_{\text{max}} - M_{\text{BH},1,\text{min}}$ panel from the left and its intersection with an $A_{\text{GWB}} = 2.8 \times 10^{-15}$ isosurface (gray).

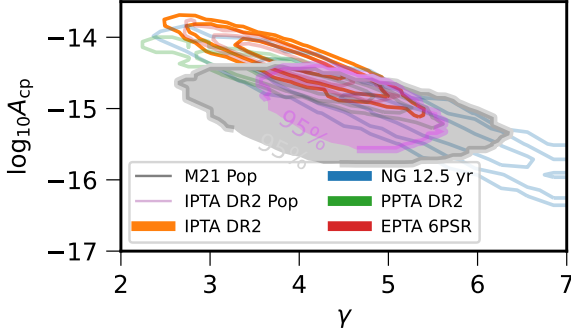


Figure 12. Comparison of power law constraints versus theoretical SMBHB populations. The 2D amplitude, spectral index constraints of the CPs from Figure 9 are compared to the region of parameters recovered from a large number of realizations of SMBHB population simulations using astrophysical relations from (M21 Pop, Middleton et al. 2021) shown in grey contours and this work (IPTA DR2 Pop) shown in purples contours.

which assumes proportionality between SMBHB and quasar populations (which may be triggered by galaxy major mergers, Stemo et al. 2020) over mass and redshift. This has the effect of setting $A_{\text{GWB}} \propto \sqrt{\Phi_{\text{BHB},0}}$, so that A_{GWB} directly implies $\Phi_{\text{BHB},0}$. To check coverage of the entire signal from SMBHBs over mass and redshift, we parameterize $A_{\text{GWB}} = A_{\text{GWB}}(\Phi_{\text{BHB},0}, M_{1,\text{min}}, z_{\text{max}})$, where $M_{1,\text{min}}$ and z_{max} are the minimum primary SMBH mass and maximum redshift, respectively, in Equation 7 (Casey-Clyde et al. 2021). We plot this parameterization of the GWB compared to various strain measurements in Figure 11, including an $A_{\text{GWB}} \approx 2.8 \times 10^{-15}$ signal (bold contour, gray isosurface).

The 2D panels of Figure 11 show three representative slices from this parameter space (one along each axis), with contours denoting their intersection with isosurfaces of constant GWB signal amplitude. The 3D plot shows the bottom right 2D panel in this 3D parameter space, along with its intersection with an $A_{\text{GWB}} \approx 2.8 \times 10^{-15}$ isosurface. We find that recovery of a background amplitude like A_{CP} requires $\Phi_{\text{BHB},0} \approx 1.5 \times 10^{-5} \text{ Mpc}^{-3}$ (corresponding to the bottom right 2D panel in Figure 11), roughly an order of magnitude larger than the $\sim 1.6 \times 10^{-6} \text{ Mpc}^{-3}$ number density implied by Mingarelli et al. (2017).

Besides this new quasar-based method the standard approach to determining the local number density $\Phi_{\text{BHB},0}$ is to model $d^3\Phi_{\text{BHB}}/(dM_1 dz dq)$ using major mergers and empirically observed galaxy and black hole relations (e.g., Simon & Burke-Spolaor 2016; Chen et al. 2019; Middleton et al. 2021). Following the methods from Middleton et al. (2021) we analyze the IPTA DR2 CP amplitude. Figure 12 compares the spread of amplitude and spectral index from the IPTA DR2 CP against values recovered from realizations of SMBHB population simulations. The original population constraints using the NANOGrav (Arzoumanian et al. 2020) frequency bins are shown by the grey shaded area. Repeating the analysis with the frequency coverage of the IPTA DR2 gives the purple shaded contours. As we reach into lower frequencies the simulations become more constrained towards the expected spectral index $\gamma = 13/3$. Limiting the SMBHB chirp mass $\mathcal{M} > 10^{8.5} M_{\odot}$ in the integral of Equation 7 we get $\Phi_{\text{BHB},0} \approx 3.0 \times 10^{-5} \text{ Mpc}^{-3}$, which is about a factor of 20 times larger than the number density from Mingarelli et al. (2017).

5.4 Outlook for other GWBs

Although we have evidence for a single common process, whose amplitude and spectral index are consistent with predicted values from a population of SMBHBs with a number density $\Phi_0 \approx 10^{-5} \text{Mpc}^{-3}$, other sources could also be plausible, for example, primordial black holes (e.g., Vaskonen & Veermäe 2021; De Luca et al. 2021; Kohri & Terada 2021), cosmic strings (e.g., Ellis & Lewicki 2021; Blasi et al. 2021; Blanco-Pillado et al. 2021) or phase transitions (e.g., Arzoumanian et al. 2021a, and references therein). These sources could produce a GWB consistent with the CP contours from PTA data.

Pol et al. (2021) have shown that the initial confident detection of a GWB including HD correlation will place very stringent constraints on the properties of the possible source of GWBs. It is also possible to have several backgrounds affecting the data, splitting the total common power into several components. A detailed study into how well we can separate multi-component GWBs is underway.

5.5 Modern Noise Mitigation

The PTA community continues to develop new data analysis strategies towards the detection of gravitational waves in pulsar timing data. Aggarwal et al. (2019, 2020) showed that unmodeled noise features in a single pulsar could leak into the gravitational wave channel for deterministic continuous GW and GW memory signals respectively. More recently immense effort was put into the development of individualized noise models for PPTA pulsars, demonstrating that sensitivity can be gained from better modeling (Goncharov et al. 2021b). Similar advanced noise modeling efforts are currently underway in both NANOGrav (Arzoumanian et al., in prep.) and EPTA (Chalumeau et al., in prep.). More sophisticated noise modeling is important, because many types of noise can add steep spectral index, low frequency power to pulsar data sets, complicating GWB recovery. For example, noise from fluctuations due to the interstellar medium or time-correlated noise from long-term instrumental effects in telescope systems, which could arise from a combination of polarisation miscalibration and secular changes in receiver gain.

6 CONCLUSION

This work shows the immense potential of combining the global efforts of PTA collaborations into one data set. Figure 1 and Figure 8 show that IPTA DR2 is significantly more sensitive than any of the constituent data sets from which it is constructed. While the data in DR2 have now been superseded by more up-to-date efforts, (Alam et al. 2021; Kerr et al. 2020; Chen et al. 2021b), Figure 9 shows that the sensitivity from combining these older data sets is comparable with these newer single PTA data releases.

The conclusions of this analysis are broadly similar to the various GWB analyses carried out by the NANOGrav, the PPTA and EPTA (Arzoumanian et al. 2020; Goncharov et al. 2021a; Chen et al. 2021b). All of these data sets favor the CP model over one with only intrinsic red noise in the individual pulsars. None of these data sets shows clear support

for the spatial correlations indicative of a GWB. Therefore a detection of a GWB can not be claimed. The strong detection of red noise that broadly matches the spectral characteristic of a GWB from SMBHBs *before* there is support for spatial correlations is expected from our understanding of the change in sensitivity of PTAs (Romano et al. 2021; Siemens et al. 2013). As was shown in (Pol et al. 2021), if the power in the auto-correlations of these pulsars is the first sign of the GWB then evidence for the spatial correlations should follow in upcoming data sets. If the next individual PTA data sets show increased support, but are short of detection thresholds, then combining them into an IPTA data set could immediately result in a data set with a significant detection. Such a combination will have the longest timespan, largest number of pulsars and independent observing systems, and will thus enable a robust GWB search.

ACKNOWLEDGEMENTS

Author contributions

An alphabetical-order author list was used for this paper to recognize the large amount of work and efforts contributed by many people within the IPTA consortium. All authors have contributed to the work via data creation, combination or analysis or contributions to the paper. The data set was created by a team led by BBPP, MED, PBD, MKE, LL, DJN, SO, SMR and MJK, by combining data sets from the constituent PTAs (EPTA, NANOGrav and PPTA). The data analysis has been the task of the Gravitational Wave Analysis Working Group (WG) of the IPTA. Many analyses have been performed over more than 2 years by PTB, AC, SChe, JAE, JMG, JSH, KI, ADJ, WGL, CMFM, NSP, NKP, DJR, LSc, JS, LSp, SRT and SJV under the leadership of several WG chairs over this time period: PTB, SChe, JSH, PDL, CMFM, DJR and SRT.

Initial tests, checks of the data quality and exploratory analyses were made by PTB, SChe, JAE, JMG, JSH, KI, CMFM, DJR, JS, SRT and SJV. Although not much of the work is presented in this paper, it helped us tremendously in understanding the subtle details of the data set and built confidence in our analysis. The final results shown come from runs performed by AC, SChe, JSH, ADJ, WGL, NSP, NKP, ASa, LSc, GS, LSp and SRT. The single pulsar noise analysis was done by SChe. The Bayesian parameter estimation were done by NSP, JSH and WGL. The optimal statistic analysis was run by JSH and NSP, with help from AC, NKP and LSp for the phase-shifts and sky-scramble null-distributions. ADJ and JSH have contributed to the Bayes factor computations. The factorized likelihood method has been applied to the data by LSc and SRT. NSP has computed the dropout factors. SChe performed the analyses of the constituent data sets, with cross-checks from ASa and GS. PTB lead the comparison between the IPTA DR2 and recent PTA data sets, which were graciously provided by JS for the NANOGrav 12.5yr, BG for the PPTA DR2 and SChe for the EPTA DR2. The astrophysical interpretation was lead by JACC and CMFM with contributions from SChe and ASa.

SChe lead the coordination of the paper writing with the help of JSH and PTB. PTB, JACC, SChe, MED, BG, JSH, CMFM, BBPP, NSP, LSc, RMS, SRT and SJV con-

tributed to the writing of the paper. The figures were created by NSP, JSH, JACC, and SChe.

Acknowledgements

The International Pulsar Timing Array (IPTA) is a consortium of existing Pulsar Timing Array collaborations, namely, the European Pulsar Timing Array (EPTA), North American Nanohertz Observatory for Gravitational Waves (NANOGrav), Parkes Pulsar Timing Array (PPTA) and the recent addition of the Indian Pulsar Timing Array (InPTA). Observing collaborations from China and South Africa are also part of the IPTA.

The EPTA is a collaboration between European and partner institutes 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. 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 NANOGrav Physics Frontiers Center is supported by NSF award number 1430284. The Green Bank Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. The Arecibo Observatory is a facility of the National Science Foundation operated under cooperative agreement by the University of Central Florida in alliance with Yang Enterprises, Inc. and Universidad Metropolitana.

The Parkes radio telescope (Murriyang) is part of the Australia Telescope which is funded by the Commonwealth Government for operation as a National Facility managed by CSIRO.

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). SBS acknowledges generous support by the NSF through grant AST-1815664. The work is supported by National SKA program of

China 2020SKA0120100, Max-Planck Partner Group, NSFC 11690024, CAS Cultivation Project for FAST Scientific. JACC was supported in part by NASA CT Space Grant PTE Federal Award Number 80NSSC20M0129. CMFM and JACC are also supported by the National Science Foundation’s NANOGrav Physics Frontier Center, Award Number 2020265. AC acknowledges support from the Paris Île-de-France Region. Support for HTC was provided by NASA through the NASA Hubble Fellowship Program grant HST-HF2-51453.001 awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. SD is the recipient of an Australian Research Council Discovery Early Career Award (DE210101738) funded by the Australian Government. GD, RK and MKr acknowledge support from European Research Council (ERC) Synergy Grant “BlackHoleCam” Grant Agreement Number 610058 and ERC Advanced Grant “LEAP” Grant Agreement Number 337062. TD is supported by the NSF AAG award number 2009468. ECF is supported by NASA under award number 80GSFC17M0002.002. BG is supported by the Italian Ministry of Education, University and Research within the PRIN 2017 Research Program Framework, n. 2017SYRTCN. Portions of this work performed at the Naval Research Laboratory is supported by NASA and ONR 6.1 basic research funding. MTL acknowledges support received from NSF AAG award number 200968. Part of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. JWMK 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). ASa, ASe and GS acknowledge financial support provided under the European Union’s H2020 ERC Consolidator Grant “Binary Massive Black Hole Astrophysics” (B Massive, Grant Agreement: 818691). RMS acknowledges support through Australian Research Council Future Fellowship FT190100155. JS acknowledges support from the JPL R&TD program. This research was funded partially by the Australian Government through the Australian Research Council (ARC), grants CE170100004 (OzGrav) and FL150100148. Pulsar research at UBC is supported by an NSERC Discovery Grant and by the Canadian Institute for Advanced Research. SRT acknowledges support from NSF grants AST-2007993 and PHY-2020265. SRT also acknowledges support from the Vanderbilt University College of Arts & Science Dean’s Faculty Fellowship program. 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).

DATA AVAILABILITY

The timing data used can be found on <https://gitlab.com/IPTA/DR2>. Selected chain files for the analyses presented in this paper can be found on <https://ipta4gw.org> and <https://doi.org/10.5281/zenodo.5787557>.

REFERENCES

- Abbott B. P., et al., 2019, *Physical Review X*, **9**, 031040
- Abbott R., et al., 2021, *Physical Review X*, **11**, 021053
- Aggarwal K., et al., 2019, *ApJ*, **880**, 116
- Aggarwal K., et al., 2020, *ApJ*, **889**, 38
- Alam M. F., et al., 2021, *ApJS*, **252**, 4
- Anholm M., Ballmer S., Creighton J. D. E., Price L. R., Siemens X., 2009, *Phys. Rev. D*, **79**, 084030
- Arzoumanian Z., et al., 2015, *ApJ*, **813**, 65
- Arzoumanian Z., et al., 2016, *ApJ*, **821**, 13
- Arzoumanian Z., et al., 2020, *ApJ*, **905**, L34
- Arzoumanian Z., et al., 2021a, arXiv e-prints, p. [arXiv:2104.13930](https://arxiv.org/abs/2104.13930)
- Arzoumanian Z., et al., 2021b, arXiv e-prints, p. [arXiv:2109.14706](https://arxiv.org/abs/2109.14706)
- Bailes M., et al., 2020, *Publ. Astron. Soc. Australia*, **37**, e028
- Blanco-Pillado J. J., Olum K. D., Wachter J. M., 2021, *Phys. Rev. D*, **103**, 103512
- Blasi S., Brdar V., Schmitz K., 2021, *Phys. Rev. Lett.*, **126**, 041305
- Burke-Spolaor S., et al., 2019, *A&ARv*, **27**, 5
- Caballero R. N., et al., 2016, *MNRAS*, **457**, 4421
- Caballero R. N., et al., 2018, *MNRAS*, **481**, 5501
- Carlin B. P., Chib S., 1995, *Journal of the Royal Statistical Society: Series B (Methodological)*, **57**, 473
- Casey-Clyde J. A., Mingarelli C. M. F., Greene J. E., Pardo K., Nañez M., Goulding A. D., 2021, in press, arXiv e-prints, p. [arXiv:2107.11390](https://arxiv.org/abs/2107.11390)
- Chamberlin S. J., Creighton J. D. E., Siemens X., Demorest P., Ellis J., Price L. R., Romano J. D., 2015, *Phys. Rev. D*, **91**, 044048
- Chen S., Sesana A., Conselice C. J., 2019, *MNRAS*, **488**, 401
- Chen Z.-C., Wu Y.-M., Huang Q.-G., 2021a, arXiv e-prints, p. [arXiv:2109.00296](https://arxiv.org/abs/2109.00296)
- Chen S., et al., 2021b, *MNRAS*, **508**, 4970
- Cornish N. J., Sampson L., 2016, *Phys. Rev. D*, **93**, 104047
- De Luca V., Franciolini G., Riotto A., 2021, *Phys. Rev. Lett.*, **126**, 041303
- Demorest P. B., et al., 2013, *ApJ*, **762**, 94
- Desvignes G., et al., 2016, *MNRAS*, **458**, 3341
- Detweiler S., 1979, *ApJ*, **234**, 1100
- Efron B., Tibshirani R. J., 1994, *An introduction to the bootstrap*. CRC press
- Ellis J., Lewicki M., 2021, *Phys. Rev. Lett.*, **126**, 041304
- Ellis J., van Haasteren R., 2017, jellis18/PTMCMCSampler: Official Release, [doi:10.5281/zenodo.1037579](https://doi.org/10.5281/zenodo.1037579), <https://doi.org/10.5281/zenodo.1037579>
- Ellis J. A., Vallisneri M., Taylor S. R., Baker P. T., 2020, ENTERPRISE: Enhanced Numerical Toolbox Enabling a Robust Pulsar Inference Suite, Zenodo, [doi:10.5281/zenodo.4059815](https://doi.org/10.5281/zenodo.4059815), <https://doi.org/10.5281/zenodo.4059815>
- Foster R. S., Backer D. C., 1990, *ApJ*, **361**, 300
- Goncharov B., et al., 2021a, arXiv e-prints, p. [arXiv:2107.12112](https://arxiv.org/abs/2107.12112)
- Goncharov B., et al., 2021b, *Mon. Not. Roy. Astron. Soc.*, **502**, 478
- Grishchuk L. P., 2005, *Physics Uspekhi*, **48**, 1235
- Hazboun J., Romano J., Smith T., 2019a, *The Journal of Open Source Software*, **4**, 1775
- Hazboun J. S., Romano J. D., Smith T. L., 2019b, *Phys. Rev. D*, **100**, 104028
- Hazboun J. S., et al., 2020a, *ApJ*, **890**, 108
- Hazboun J. S., Simon J., Siemens X., Romano J. D., 2020b, *ApJ*, **905**, L6
- Hee S., Handley W. J., Hobson M. P., Lasenby A. N., 2016, *MNRAS*, **455**, 2461
- Hellings R. W., Downs G. S., 1983, *ApJ*, **265**, L39
- Hobbs G., et al., 2010, *Classical and Quantum Gravity*, **27**, 084013
- Hobbs G., et al., 2020, *MNRAS*, **491**, 5951
- Jaffe A. H., Backer D. C., 2003, *ApJ*, **583**, 616
- Joshi B. C., et al., 2018, *Journal of Astrophysics and Astronomy*, **39**, 51
- Kaspi V. M., Taylor J. H., Ryba M., 1994, *ApJ*, **428**, 713
- Keith M. J., et al., 2013, *MNRAS*, **429**, 2161
- Kelley L. Z., Blecha L., Hernquist L., Sesana A., Taylor S. R., 2017, *MNRAS*, **471**, 4508
- Kerr M., et al., 2020, *Publ. Astron. Soc. Australia*, **37**, e020
- Kohri K., Terada T., 2021, *Physics Letters B*, **813**, 136040
- Lam M. T., et al., 2018, *ApJ*, **861**, 132
- Lasky P. D., et al., 2016, *Physical Review X*, **6**, 011035
- Lee K. J., 2016, in Qian L., Li D., eds, *Astronomical Society of the Pacific Conference Series Vol. 502, Frontiers in Radio Astronomy and FAST Early Sciences Symposium 2015*. p. 19
- Lentati L., et al., 2015, *MNRAS*, **453**, 2576
- Lentati L., et al., 2016, *MNRAS*, **458**, 2161
- Mahalanobis P. C., 1936, in *On the Generalized Distance in Statistics*. Proceedings of the National Institute of Science of India, pp 49–55
- Manchester R. N., et al., 2013, *Publ. Astron. Soc. Australia*, **30**, e017
- Middleton H., Sesana A., Chen S., Vecchio A., Del Pozzo W., Rosado P. A., 2021, *MNRAS*, **502**, L99
- Mingarelli C. M. F., et al., 2017, *Nat. Astron.*, **1**, 886
- Ölmez S., Mandic V., Siemens X., 2010, *Phys. Rev. D*, **81**, 104028
- Perera B. B. P., et al., 2019, *MNRAS*, **490**, 4666
- Phinney E. S., 2001, arXiv e-prints, [pp astro-ph/0108028](https://arxiv.org/abs/astro-ph/0108028)
- Pol N. S., et al., 2021, *Astrophys. J. Lett.*, **911**, L34
- Rajagopal M., Romani R. W., 1995, *ApJ*, **446**, 543
- Reardon D. J., et al., 2016, *MNRAS*, **455**, 1751
- Roebber E., 2019, *ApJ*, **876**, 55
- Roebber E., Holder G., 2017, *ApJ*, **835**, 21
- Romano J. D., Hazboun J. S., Siemens X., Archibald A. M., 2021, *Phys. Rev. D*, **103**, 063027
- Sazhin M. V., 1978, *Soviet Ast.*, **22**, 36
- Sesana A., 2013a, *Classical and Quantum Gravity*, **30**, 224014
- Sesana A., 2013b, *MNRAS*, **433**, L1
- Sesana A., Vecchio A., Colacino C. N., 2008, *MNRAS*, **390**, 192
- Shannon R. M., Cordes J. M., 2010, *ApJ*, **725**, 1607
- Shannon R. M., et al., 2015, *Science*, **349**, 1522
- Siemens X., Ellis J., Jenet F., Romano J. D., 2013, *Classical and Quantum Gravity*, **30**, 224015
- Simon J., Burke-Spolaor S., 2016, *ApJ*, **826**, 11
- Stemo A., Comerford J. M., Barrows R. S., Stern D., Assef R. J., Griffith R. L., Schechter A., 2020, arXiv e-prints, p. [arXiv:2011.10051](https://arxiv.org/abs/2011.10051)
- Taylor S. R., Lentati L., Babak S., Brem P., Gair J. R., Sesana A., Vecchio A., 2017, *Phys. Rev. D*, **95**, 042002
- Tiburzi C., et al., 2016, *MNRAS*, **455**, 4339
- Vallisneri M., et al., 2020, *ApJ*, **893**, 112
- Vaskonen V., Veermäe H., 2021, *Phys. Rev. Lett.*, **126**, 051303
- Verbiest J. P. W., et al., 2016, *MNRAS*, **458**, 1267
- Vigeland S. J., Islo K., Taylor S. R., Ellis J. A., 2018, *Phys. Rev. D*, **98**, 044003
- Zhu W. W., et al., 2015, *ApJ*, **809**, 41
- van Haasteren R., Levin Y., McDonald P., Lu T., 2009, *MNRAS*, **395**, 1005

¹ Institute of Astrophysics, FORTH, N. Plastira 100, 70013, Heraklion, Greece

² Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, 53121 Bonn, Germany

³ Argelander Institut für Astronomie, Auf dem Hügel 71, 53117, Bonn, Germany

⁴ X-Ray Astrophysics Laboratory, NASA Goddard Space Flight Center, Code 662, Greenbelt, MD 20771, USA

⁵ Université de Paris, CNRS, Astroparticule et Cosmologie, 75013 Paris, France

⁶ Moscow Institute of Physics and Technology, Dolgo-

prudny, Moscow region, Russia

⁷ Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, VIC 3122, Australia

⁸ Australian Research Council Centre for Excellence for Gravitational Wave Discovery (OzGrav)

⁹ Fakultät für Physik, Universität Bielefeld, Postfach 100131, 33501 Bielefeld, Germany

¹⁰ Department of Physics and Astronomy, Widener University, One University Place, Chester, PA 19013, USA

¹¹ ASTRON, Netherlands Institute for Radio Astronomy, Oude Hoogeveensedijk 4, 7991 PD, Dwingeloo, The Netherlands

¹² Department of Physics, Montana State University, Bozeman, MT 59717, USA

¹³ Laboratoire de Physique et Chimie de l'Environnement et de l'Espace LPC2E UMR7328, Université d'Orléans, CNRS, 45071 Orléans, France

¹⁴ Station de Radioastronomie de Nançay, Observatoire de Paris, PSL University, CNRS, Université d'Orléans, 18330 Nançay, France

¹⁵ Dipartimento di Fisica "G. Occhialini", Università degli Studi di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy

¹⁶ INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, 20126 Milano, Italy

¹⁷ Cornell Center for Astrophysics and Planetary Science and Department of Astronomy, Cornell University, Ithaca, NY 14853, USA

¹⁸ Department of Physics and Astronomy, West Virginia University, P.O. Box 6315, Morgantown, WV 26506, USA

¹⁹ Center for Gravitational Waves and Cosmology, West Virginia University, Chestnut Ridge Research Building, Morgantown, WV 26505, USA

²⁰ INAF - Osservatorio Astronomico di Cagliari, via della Scienza 5, 09047 Selargius (CA), Italy

²¹ Canadian Institute for Advanced Research, CIFAR Azrieli Global Scholar, MaRS Centre West Tower, 661 University Ave. Suite 505, Toronto ON M5G 1M1, Canada

²² Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, P. R. China

²³ Department of Physics, University of Connecticut, 196 Auditorium Road, U-3046, Storrs, CT 06269-3046, USA

²⁴ Department of Physics and Astronomy, Vanderbilt University, 2301 Vanderbilt Place, Nashville, TN 37235, USA

²⁵ Department of Physics and Astronomy, Franklin & Marshall College, P.O. Box 3003, Lancaster, PA 17604, USA

²⁶ NASA Hubble Fellowship Program Einstein Postdoctoral Fellow

²⁷ Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada

²⁸ School of Science, Western Sydney University, Locked Bag 1797, Penrith South DC, NSW 2751, Australia

²⁹ George Mason University, resident at the Naval Research Laboratory, Washington, DC 20375, USA

³⁰ National Radio Astronomy Observatory, 1003 Lopezville Rd., Socorro, NM 87801, USA

³¹ LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de Paris, 5 place Jules Janssen, 92195 Meudon, France

³² Department of Physics, Hillsdale College, 33 E. College Street, Hillsdale, MI 49242, USA

³³ Eureka Scientific, Inc., 2452 Delmer Street, Suite 100, Oakland, CA 94602-3017, USA

³⁴ School of Physics and Astronomy, Rochester Institute of Technology, Rochester, NY 14623, USA

³⁵ Department of Astronomy, University of Maryland, College Park, MD 20742, USA

³⁶ Center for Research and Exploration in Space Science and Technology, NASA/GSFC, Greenbelt, MD 20771, USA

³⁷ NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

³⁸ Department of Physics and Astronomy, West Virginia University, PO Box 6315, Morgantown, WV 26506, USA

³⁹ Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Am Mühlberg 1, 14476 Potsdam, Germany

⁴⁰ Gran Sasso Science Institute (GSSI), 67100 L'Aquila, Italy

⁴¹ INFN, Laboratori Nazionali del Gran Sasso, 67100 Assergi, Italy

⁴² Physical Sciences Division, University of Washington Bothell, 18115 Campus Way NE, Bothell, WA 98011, USA

⁴³ Australia Telescope National Facility, CSIRO Space and Astronomy, P.O. Box 76, Epping, NSW 1710, Australia

⁴⁴ Center for Gravitation, Cosmology and Astrophysics, Department of Physics, University of Wisconsin-Milwaukee, P.O. Box 413, Milwaukee, WI 53201, USA

⁴⁵ Department of Astrophysics/IMAPP, Radboud University Nijmegen, P.O. Box 9010, 6500 GL Nijmegen, The Netherlands

⁴⁶ Jodrell Bank Centre for Astrophysics, Department of Physics and Astronomy, University of Manchester, Manchester M13 9PL, UK

⁴⁷ Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA), Northwestern University, Evanston, IL 60208, USA

⁴⁸ Space Science Division, Naval Research Laboratory, Washington, DC 20375, USA

⁴⁹ Laboratory for Multiwavelength Astrophysics, Rochester Institute of Technology, Rochester, NY 14623, USA

⁵⁰ Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA

⁵¹ National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, P. R. China

⁵² Astrophysics Group, Cavendish Laboratory, JJ Thomson Avenue, Cambridge, CB3 0HE, UK

⁵³ Department of Astronomy & Astrophysics, University of Toronto, 50 Saint George Street, Toronto, ON M5S 3H4, Canada

⁵⁴ Green Bank Observatory, P.O. Box 2, Green Bank, WV 24944, USA

⁵⁵ Department of Physics, University of the Pacific, 3601 Pacific Ave., Stockton, CA 95211, USA

⁵⁶ Canadian Institute for Theoretical Astrophysics, University of Toronto, 60 St. George Street, Toronto, ON M5S 3H8, Canada

⁵⁷ Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA

⁵⁸ Dunlap Institute for Astronomy and Astrophysics, University of Toronto, 50 Saint George Street, Toronto, ON M5S 3H4, Canada

⁵⁹ Department of Physics, Lafayette College, Easton, PA 18042, USA

⁶⁰ Manly Astrophysics, 15/41-42 East Esplanade, Manly, NSW 2095, Australia

⁶¹ Institute of Physics, Eötvös Loránd University, Pázmány P.s. 1/A, 1117 Budapest, Hungary

⁶² Arecibo Observatory, HC3 Box 53995, Arecibo, PR 00612, USA

⁶³ Università di Cagliari, Dipartimento di Fisica, S.P. Monserrato-Sestu Km 0,700 - 09042 Monserrato (CA), Italy

⁶⁴ National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, VA 22903, USA

⁶⁵ U.S. Naval Research Laboratory, Washington, DC 20375, USA

⁶⁶ CSIRO Scientific Computing Services, Australian Technology Park, Locked Bag 9013, Alexandria, NSW 1435, Australia

⁶⁷ Australia Telescope National Facility, CSIRO Space and Astronomy, P.O. Box 276, Parkes, NSW 2870, Australia

⁶⁸ Theoretical Physics Department, CERN, 1211 Geneva 23, Switzerland

⁶⁹ Department of Physics, Oregon State University, Corvallis, OR 97331, USA

⁷⁰ Department of Astrophysical and Planetary Sciences, University of Colorado, Boulder, CO 80309, USA

⁷¹ Department of Physics and Astronomy, Swarthmore College, Swarthmore, PA 19081, USA

⁷² Department of Physics and Astronomy, Oberlin College, Oberlin, OH 44074, USA

⁷³ Laboratoire Univers et Théories LUTH, Observatoire de Paris, Université PSL, CNRS, Université de Paris, 92190 Meudon, France

⁷⁴ Theoretical Astrophysics Including Relativity (TAPIR), MC 350-17, California Institute of Technology, Pasadena, CA 91125, USA

⁷⁵ Institute for Gravitational Wave Astronomy and School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

⁷⁶ Xinjiang Astronomical Observatory, Chinese Academy of Sciences, 150 Science 1-Street, Urumqi, Xinjiang 830011, P. R. China

⁷⁷ Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing 210023, P. R. China

⁷⁸ Advanced Institute of Natural Sciences, Beijing Normal University at Zhuhai 519087, P. R. China