

# Rapid construction of joint pulsar timing array data sets: the Lite method

B. Larsen <sup>1</sup>  C. M. F. Mingarelli <sup>1</sup>  P. T. Baker <sup>1</sup>  J. S. Hazboun <sup>1</sup>  S. Chen <sup>1</sup>  L. Schult <sup>1</sup>  S. R. Taylor <sup>1</sup>  J. Simon <sup>1</sup>  J. Antoniadis <sup>1</sup>  J. Baier <sup>1</sup>  R. N. Caballero <sup>1</sup>  A. Chalumeau <sup>1</sup>  Z. Chen <sup>1</sup>  I. Cognard <sup>1</sup>  D. Deb <sup>1</sup>  V. Di Marco <sup>1</sup>  T. Dolch <sup>1</sup>  I. O. Eya <sup>1</sup>  E. C. Ferrara <sup>1</sup>  K. A. Gersbach <sup>1</sup>  D. C. Good <sup>1</sup>  H. Hu <sup>1</sup>  A. Kapur <sup>1</sup>  S. Kala <sup>1</sup>  M. Kramer <sup>1</sup>  M. T. Lam <sup>1</sup>  W. G. Lamb <sup>1</sup>  T. J. W. Lazio <sup>1</sup>  K. Liu <sup>1</sup>  Y. Liu <sup>1</sup>  M. McLaughlin <sup>1</sup>  D. J. Nice <sup>1</sup>  B. B. P. Perera <sup>1</sup>  A. Petiteau <sup>1</sup>  S. M. Ransom <sup>1</sup>  D. J. Reardon <sup>1</sup>  C. J. Russell <sup>1</sup>  G. M. Shaifullah <sup>1</sup>  L. Speri <sup>1</sup>  A. Srivastava <sup>1</sup>  G. Theureau <sup>1</sup>  J. Wang <sup>1</sup>  J. Wang <sup>1</sup>  and L. Zhang <sup>1</sup> 

*Affiliations are listed at the end of the paper*

Accepted 2025 August 26. Received 2025 August 25; in original form 2025 March 26

## ABSTRACT

The International Pulsar Timing Array (IPTA)’s second data release (IPTA DR2) combines decades of observations of 65 millisecond pulsars from 7 radio telescopes. IPTA data sets should be the most sensitive data sets to nanohertz gravitational waves (GWs), but take years to assemble, often excluding valuable recent data. To address this, we introduce the IPTA ‘Lite’ analysis, where a Figure of Merit is used to select an optimal PTA data set to analyse for each pulsar, enabling immediate access to new data and preliminary results prior to full combination. We test the capabilities of the Lite analysis using IPTA DR2, finding that ‘DR2 Lite’ can be used to detect the common red noise process with an amplitude of  $A = 4.8_{-1.8}^{+1.8} \times 10^{-15}$  at  $\gamma = 13/3$ . This amplitude is slightly large in comparison to the combined analysis, and likely biased high as DR2 Lite is more sensitive to systematic errors from individual pulsars than the full data set. Furthermore, although there is no strong evidence for Hellings-Downs correlations in IPTA DR2, we still find the full data set is better at resolving Hellings-Downs correlations than DR2 Lite. Alongside the Lite analysis, we also find that analysing a subset of pulsars from IPTA DR2, available at a hypothetical ‘early’ stage of combination (EDR2), yields equally competitive results as the full data set. Looking ahead, the Lite method will enable rapid synthesis of the latest PTA data, offering preliminary GW constraints before the superior full data set combinations are available.

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

## 1 INTRODUCTION

Pulsar Timing Arrays (PTAs) are experiments for detecting low-frequency gravitational waves (GWs), offering unprecedented access to nanohertz GWs (Sazhin 1978; Detweiler 1979; Hellings & Downs 1983). By monitoring the ultra-stable arrival times of radio pulses from millisecond pulsars (MSPs) – nature’s most precise clocks – we can build a galaxy-scale GW detector. PTAs use deviations in the time of arrivals (TOAs) induced by GWs to infer their presence, enabling the detection of signals from supermassive black hole binaries (SMBHBs; Begelman, Blandford & Rees 1980), the stochastic gravitational wave background (GWB) that should arise from their cosmic merger history (Rajagopal & Romani 1995; Volonteri, Haardt & Madau 2003), and potentially new physics (Lasky et al. 2016; Caprini & Figueroa 2018; Afzal et al. 2023).

Currently, PTA collaborations around the globe, including the European PTA (EPTA; Desvignes et al. 2016), the North Amer-

ican Nanohertz Observatory for Gravitational Waves (NANOGrav; Ransom et al. 2019), the Parkes PTA (PPTA; Manchester et al. 2013), the Chinese PTA (CPTA; Lee 2016), the Indian PTA (InPTA; Joshi et al. 2018), the MeerKAT PTA (MPTA; Miles et al. 2023), and the Fermi  $\gamma$ -ray PTA ( $\gamma$ PTA; FERMI-LAT Collaboration 2022) provide PTA data sets of varying sensitivity and duration. The International Pulsar Timing Array (IPTA) is a consortium of PTA collaborations, with the first and second data releases, IPTA DR1 (Verbiest et al. 2016) and IPTA DR2 (Perera et al. 2019), each combining pulsar observations from EPTA, NANOGrav, and PPTA into one coherent data set. A third IPTA data release, IPTA DR3, is currently in development, with the potential to include data from all current PTAs.

Data combination improves a PTA’s sensitivity to GWs by increasing the pulsars’ effective observation timespan, cadence, and radio frequency coverage, as well as the PTA’s overall number of pulsars and sky coverage. This was expected by Siemens et al. (2013) and verified in the IPTA DR2 GWB search (Antoniadis et al. 2022), which found evidence for a common red noise (CRN; also known as CURN, Chen et al. 2021a; Agazie et al. 2024) process, a signature of an emerging GWB signal, with greater significance

\* E-mail: [bjorn.larsen@yale.edu](mailto:bjorn.larsen@yale.edu)

**Table 1.** Names and short descriptions of the three data sets we analyse in this work. Section 2 provides further details on each of the data sets.

Name	Short description
Full DR2	The fully-combined IPTA DR2 from Perera et al. (2019), with $T_{\text{obs}} > 3$ yr filter from Antoniadis et al. (2022) for 53 pulsars in total.
DR2 Lite	The Lite data set presented in this work, with a total of 53 pulsars selected using the FoM from single-PTA data subsets of Full DR2.
EDR2	An ‘early’ subset of Full DR2 which includes fully-combined data for only 22 pulsars corresponding to the highest FoM in DR2 Lite.

than the individual constituent data sets comprising IPTA DR2. This provided a robust confirmation of the CRN first found by the regional PTA collaborations (Arzoumanian et al. 2020b; Chen et al. 2021b; Goncharov et al. 2021b).

While it did not produce the first published measurement of a CRN process, we point out that IPTA DR2 could have presented the earliest opportunity to detect the CRN, as regional PTAs needed to collect approximately  $\sim 2$  more years of data than was used for IPTA DR2 in order to sufficiently resolve the signal. The delay in the IPTA DR2 analysis largely resulted from the resource-intensive process of data combination, which requires meticulous handling of different kinds of data, alignment of timing models, fitting for instrumental offsets, and iterative noise modelling across data sets (Verbiest et al. 2016; Perera et al. 2019). This effort typically takes several years to complete, meaning by the time a combined data set is released, portions of the underlying data are already outdated. To illustrate the time-scales, IPTA DR2 contains the NANOGrav 9-yr data set (Arzoumanian et al. 2016), but IPTA DR2 was not published until 3 yr later (Perera et al. 2019). By the time the IPTA DR2 GWB search was carried out (Antoniadis et al. 2022), the NANOGrav 12.5-yr data set was already used to detect the CRN (Arzoumanian et al. 2020a). Thus, the CRN could theoretically have been measured 3 yr in advance of Arzoumanian et al. (2020a) if IPTA DR2 was constructed immediately. Given the long time-scales required for new GW signals to emerge in PTA data sets, this motivates the need to either improve the speed of data combination or explore alternative methods for analysing joint-PTA data sets, supplementing the eventual results of a fully-combined data set.

To address this, we introduce a novel, resource-efficient approach we call the ‘Lite’ method. Instead of immediately performing full data combination, we evaluate pulsar data sets which have already been produced by individual PTAs, and select the most informative data set for each pulsar using a Figure of Merit (FoM), which quantifies sensitivity to a GW signal based on the data set properties using the theoretical scaling laws for the signal-to-noise ratio (S/N). For example, the GWB S/N from Siemens et al. (2013) suggests a FoM which increases as total observing time increases, as observation cadence increases, and as RMS white noise residual decreases. By selecting each pulsar’s data based on the FoM, the Lite data set achieves the maximum theoretical sensitivity to GW signals possible among all available data prior to performing data combination. As such, a Lite data set analysis may provide an early look into what may result from a fully-combined analysis. A Lite data set will also be less computationally intensive to analyse than its fully-combined counterpart due to the reduced data volume.

Data combination is a slow, intensive process, with combined data sets built up one pulsar at a time. It is tempting then to also consider the result of analysing an *early* or *intermediate* combined data set, which includes just the first set of pulsars which have had their data combined. The FoM suggests which pulsars to combine first: combining pulsars in order, starting from highest FoM to lowest, will maximize the GWB sensitivity of any intermediate combined data set. This practice already has precedent within PTA analyses (Babak et al. 2016; Speri et al. 2023). For example, the creation of EPTA

DR2 with 25 pulsars (EPTA Collaboration 2023a) was preceded by a version of EPTA DR2 using 6 pulsars (Chen et al. 2021c), which were originally selected based on their expected S/N for continuous GWs (Babak et al. 2016). The 25 pulsars used for the EPTA DR2 GWB search (EPTA Collaboration 2023a, c) were then selected to optimize the theoretical S/N of the Hellings-Downs curve, following the method from Speri et al. (2023).

Here, we use IPTA DR2 as a test case to assess the benefits of performing GW searches using a *Lite* data set, an *early-combined* data set, and a *fully-combined* data set, reflecting the stages in which future combined data may be analysed. Table 1 provides short names and descriptions of each of these data sets for ease of reference. Specifically, we test how the detection statistics and upper limits for a GWB evolve as more data are combined. This analysis framework thus quantifies the benefits and drawbacks of a rapid, on-the-fly Lite analysis, as well as the superior sensitivity offered by the full data combination.

The information from this analysis will also be valuable for the interpretation of present-day data sets. Agazie et al. (2024) performed comparisons and joint-analyses of the NANOGrav 15 yr data set (Agazie et al. 2023a, b), EPTA + InPTA DR2 (EPTA Collaboration 2023a, c), and PPTA DR3 (Reardon et al. 2023a; Zic et al. 2023a). The factorized likelihood cross-PTA analyses in Agazie et al. (2024) are similar in spirit to the Lite analysis method we present here, and the results suggest that IPTA DR3 will place the most decisive detection to date of the Hellings-Downs curve, which is the definitive signature of an isotropic GWB imprinted in the cross-correlations between pulsar timing residuals (Hellings & Downs 1983). Our Lite analysis of IPTA DR2 will therefore be useful to calibrate expectations for IPTA DR3.

Our paper is laid out as follows: In Section 2, we detail IPTA DR2, the FoM, and our method of creating DR2 Lite from IPTA DR2 as a starting point. In Section 3, we describe the PTA likelihood, models, and parameters used in our Bayesian analysis of each data set. In Section 4, we assess how the statistics for both a CRN and an Hellings-Downs cross-correlated GWB evolve throughout each stage of data combination, as well as the impact of data combination on single pulsar noise characterization and ensemble noise properties. In Section 5, we discuss our results and future directions for the Lite method.

## 2 IPTA DR2 AND THE LITE DATA SET

IPTA DR2 is the most recent IPTA-combined data set, fully detailed in Perera et al. (2019). Steps required to create the combined data set include standardization of TOA flags (metadata), fitting for instrumental offsets, implementation of comprehensive timing and noise models, and simultaneous/iterative fits to the pulsar timing and noise model parameters (Verbiest et al. 2016; Perera et al. 2019). IPTA DR2 includes TOAs from the NANOGrav 9-yr data set (Arzoumanian et al. 2016), EPTA DR1 (Desvignes et al. 2016), and PPTA DR1 (Manchester et al. 2013; Reardon et al. 2016), as well as legacy NANOGrav timing data for PSRs J1713+0747, J1857+0943, and J1939+2134 (Kaspi, Taylor & Ryba 1994; Zhu

et al. 2015) and extended PPTA data for PSRs J0437–4715, J1744–1134, J1713+0747, and J1909–3744 (Shannon et al. 2015). In total, IPTA DR2 includes a total of 65 millisecond pulsars, with data sets spanning 0.5–30 yr, measured across 7 different telescopes. IPTA DR2 also features two versions, designated VersionA and VersionB, which were each created using different noise models. Throughout this work we use only VersionB. A few pulsars in IPTA DR2 are known by their Besselian names in NANOGRAv data sets, though in this work we use their Julian names: J1857+0943 (B1855+09), J1939+2134 (B1937+21), and J1955+2908 (B1953 + 29).

Antoniadis et al. (2022) carried out a GWB search on a subset of 53 pulsars from IPTA DR2 with  $> 3$  yr of data. Pulsars with shorter timespans do not resolve the GWB at low frequencies, and their timing models may not yet be converged (Andrews, Lam & Dolch 2020). This data set, designated here as Full DR2, is our benchmark against which to compare the results of the Lite analysis. The search from Antoniadis et al. (2022) yielded a strong detection of a CRN process (the autocorrelated component of a GWB). Detection statistics for Hellings-Downs correlations (the cross-correlated component of a GWB) were also computed, but the values were considered insufficient for a detection.

IPTA DR2's very long observation timespan of 30.2 yr results from the inclusion of legacy data no longer used in some more recent PTA data releases (EPTA Collaboration 2023a; Agazie et al. 2023b; Zic et al. 2023b). Among these legacy data include TOAs which have been observed only at single radio frequencies, which are suboptimal for accurately characterizing DM variations (Shannon & Cordes 2017; Lam et al. 2018a; Sosa Fiscella et al. 2024). It has since been shown empirically with EPTA DR2 (EPTA Collaboration 2023b, c) and simulations (Ferranti et al. 2025) that including these types of single-frequency data can reduce the sensitivity of the PTA to cross-correlations between pulsar pairs, which must be used to resolve the Hellings-Downs curve. For simplicity and consistency with Antoniadis et al. (2022), we include these TOAs in all versions of our analysis, but highlight that their presence should be considered during the interpretation of our results. We reserve an analysis assessing the impacts of legacy data in IPTA data sets for future work.

## 2.1 Selecting pre-combined pulsar data using a Figure of Merit

We next detail the methods of creating an IPTA Lite data set composed of *pre*-combined data from individual PTA data sets. In the presence of multiple data sets for a given pulsar, we select whichever data maximizes a FoM. The FoM encodes the theoretical sensitivity of a pulsar to a particular GW signal, based solely on the properties of the pulsar's TOAs. For this work, we define our FoM according to the scaling laws for a GWB with spectral index  $\gamma = 13/3$  in the intermediate regime where the lowest frequencies of the GWB have risen above the white noise level (Siemens et al. 2013),

$$\text{FoM}_{\text{GWB}} = \left( \frac{T_{\text{obs}}}{\left( \langle \sigma_{\text{TOA}} \rangle^2 \langle \Delta t \rangle \right)^{3/13}} \right)^{1/2}, \quad (1)$$

where  $T_{\text{obs}}$  is the pulsar's total observation timespan,  $\langle \sigma_{\text{TOA}} \rangle$  is the average (harmonic mean) TOA error, and  $\langle \Delta t \rangle$  is the average (geometric mean) time between observations. Intuitively, equation (1) rewards data sets with high data quantity, i.e. long timespans  $T_{\text{obs}}$  and high data cadence  $c = 1/\Delta t$ , as well as data sets with high data quality, i.e. smaller errors  $\sigma_{\text{TOA}}$ . The factor of 3/13 results from the predicted spectral index of the emerging GWB. The intermediate regime GWB S/N is also proportional to the number of pulsars, but

this is not included in the FoM to select which data to use for a single pulsar. Different Lite data sets can be also curated for different nHz GW searches as the theoretical S/N for each type of GW signal will follow a different scaling law. We additionally present the FoM for continuous GWs and for GW bursts with memory in Appendix A, but we do not explore these in further in this work.

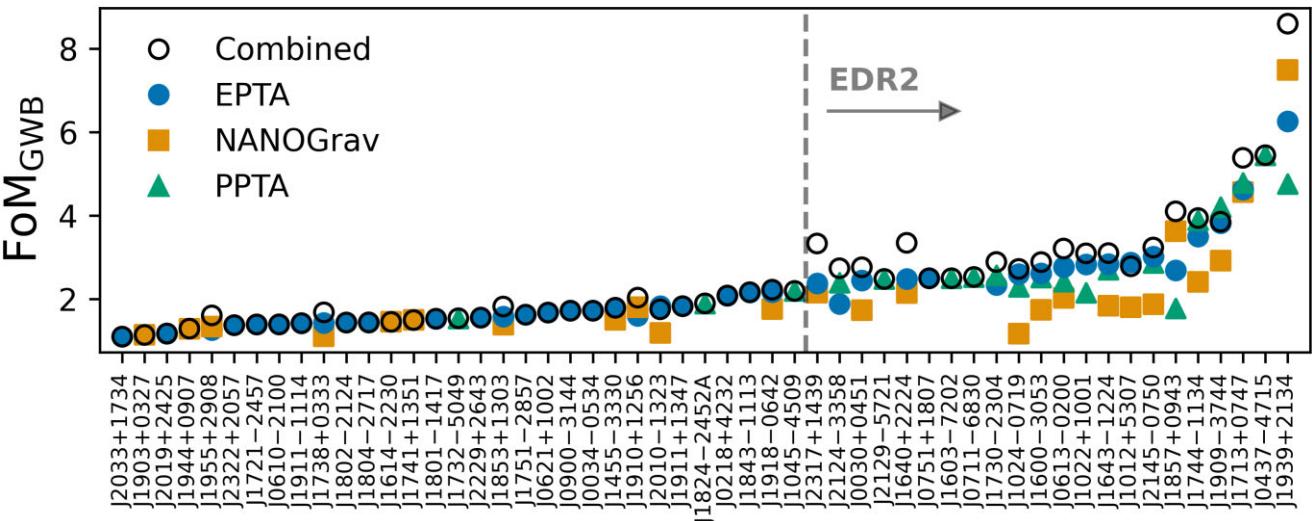
To create a Lite data set, first decide which GW signal to search for and select the FoM. Next, iterate through all pulsars of interest. Pulsars timed by a single PTA require no extra work to include them in the Lite data set, aside from ensuring terrestrial clock references and Solar system ephemeris versions are consistent and up to date among all pulsars. If a pulsar is timed by multiple PTAs, compute the FoM from each PTA's data for that pulsar, then take whichever FoM is largest and add the corresponding PTA's data to the Lite data set. The Lite data set is therefore the bespoke composition of uncombined pulsar data sets across different PTAs, which may be used from there to perform a joint GW search. The FoM-based selection approach has the advantage that it is purely based on the statistical properties of the data set itself and is agnostic to which PTA timed it. Lite data sets may be created immediately from the latest PTA data releases.

## 2.2 Creating intermediate data sets from IPTA DR2

We next detail the specifics behind the curation of our IPTA DR2-based data subsets, which together with Full DR2 are summarized Table 1. IPTA DR2 is the most recent fully-combined IPTA data set and is therefore an excellent data set to test the performance of the Lite analysis. Here, we choose to create DR2 Lite *starting* from the IPTA DR2 release. This choice ensures the timing models are identical and thus any difference in GW sensitivity between DR2 Lite and Full DR2 results purely from the difference in data volume. In principle, one should also refit the timing models after reducing the data volume, or at least check to ensure the remaining timing residuals are within the regime of the linear timing model used during GW analyses. We empirically found the latter assumption to hold for the Lite version of IPTA DR2. Additionally, a maximally sensitive 'early' version of a fully-combined data set will start with a subset of pulsars that maximize the FoM. We use IPTA DR2 to create this hypothetical 'early' data set, which we call EDR2, drawing inspiration from the *Gaia* data releases (Gaia Collaboration 2021).

To create DR2 Lite starting in IPTA DR2, we first isolate each PTA's TOAs in each pulsar and compute the FoM from equation (1). Each pulsar in DR2 Lite then keeps only the PTA data with the largest FoM. Fig. 1 visualizes the results of this process, by comparing the FoM computed for each PTA and each pulsar. DR2 Lite in total uses EPTA data for 33 pulsars, NANOGRAv data for 8 pulsars, and PPTA data for 12 pulsars. Fig. 1 also shows each pulsar's FoM computed from the fully-combined data set—in nearly all cases the combined data result in a higher FoM, as expected from equation (1). The FoM is slightly lower only for PSRs J1012+5307 and J1909–3744 using the combined data. Following equation (1), this results if the newly combined TOAs have much larger errors on average than the TOAs included in the Lite data set and the timespan and data cadence do not appreciably increase in contrast.

We also use the FoM distribution in Fig. 1 to select which pulsars to include in EDR2. Specifically, we rank the pulsars in order of highest to lowest FoM, as computed from DR2 Lite. This ranking represents an optimal order for combining the data to maximize the sensitivity of the early-combined data set to a GWB. For EDR2, we choose the 22 highest ranked pulsars, represented by all pulsars to the right of the vertical line in Fig. 1. We choose to use a 22 pulsar cut-off for this analysis for a number of reasons, though the exact



**Figure 1.** Figures of Merit (FoMs; equation 1) computed for each pulsar in IPTA DR2. The different markers indicate whether the FoM is computed either using the EPTA (blue circles), NANOGrav (orange squares), or PPTA (green triangles) TOAs, while the open black circles indicate the pulsar's FoM computed using all TOAs together. The PTA data with the largest FoM for each pulsar is used in DR2 Lite. From left to right, the pulsars are ranked in order from lowest to highest FoM using whichever PTA's data are included in the Lite data set. All pulsars to the right of the dashed vertical line are included in EDR2.

number is ultimately arbitrary. Namely, this number should be large enough to avoid bias in GWB statistics due to a finite number of pulsars (Johnson et al. 2022), but still represents less than half of the total pulsars intended for the full analysis. This also cleanly cuts off the data set at PSR J2317+1439, which sees a large boost in the FoM post-combination based on Fig. 1. Finally, this was also selected as a rough match for the number of pulsars with data combined for the upcoming IPTA DR3 at the time of performing this analysis (Good & International Pulsar Timing Array Team 2023).

Fig. 2 further visualizes each data set by displaying the observation times and radio frequencies of each TOA. Coloured markers represent the single-PTA data used in DR2 Lite, while coloured + black markers represent all data used in Full DR2. EDR2 pulsars are highlighted in gold. Several pulsars (e.g. PSR J0437–4715) are timed only by a single PTA; therefore, their Lite and combined data sets are identical. Other pulsar's combined data (e.g. PSR J1713+0747) present a clear advantage in total radio band coverage and observation cadence, especially in the latter half of the data set. For reproducibility, code for creating Lite data sets from IPTA DR2 can be found in the public IPTA github repository.<sup>1</sup>

### 3 ANALYSIS METHODS

We closely follow the methods and conventions in Antoniadis et al. (2022) for the Bayesian analysis of IPTA DR2 and its corresponding Lite data set. This will be review for readers familiar with the prior work, although in Section 3.2 we also update the pulsar noise models over those used in Antoniadis et al. (2022). We use a multivariate Gaussian likelihood to represent our timing residual vector  $\delta t$  under the full signal model  $\mathcal{M}$ , expressed compactly as

$$\mathcal{L}(\delta t | \eta, \mathcal{M}) = \frac{\exp\left(-\frac{1}{2}(\delta t - s)^T \mathbf{C}^{-1}(\delta t - s)\right)}{\sqrt{\det(2\pi \mathbf{C})}}, \quad (2)$$

<sup>1</sup>[https://github.com/ipta/IPTA\\_DR2\\_analysis/blob/master/gen\\_FoM\\_data\\_set.ipynb](https://github.com/ipta/IPTA_DR2_analysis/blob/master/gen_FoM_data_set.ipynb)

where  $\eta$  is a vector of GW signal and noise model (hyper)parameters,  $s$  is the mean model accounting for any signals we choose to model deterministically, and  $\mathbf{C}$  is the  $N_{\text{TOA}} \times N_{\text{TOA}}$  covariance matrix (van Haasteren & Vallisneri 2014). The covariance matrix is expressed  $\mathbf{C} = \mathbf{N} + \mathbf{T}\mathbf{B}\mathbf{T}^T$ , where  $\mathbf{N}$  is a block diagonal matrix accounting for TOA errors and additional white noise parameters. All remaining signals are modelled as rank-reduced Gaussian processes in  $\mathbf{T}\mathbf{B}\mathbf{T}^T$ , where  $\mathbf{T}$  is the concatenation of several  $N_{\text{TOA}} \times N_b$  design matrices of  $N_b$  basis functions, and  $\mathbf{B} = \langle \mathbf{b}\mathbf{b}^T \rangle$  encodes the variance over weights  $\mathbf{b}$ . Subvectors of  $\mathbf{b}$  belong to the weight space  $\mathcal{W}$  and submatrices of  $\mathbf{T}$  belong to the space of design matrices  $\mathcal{T}$ . The posterior distribution over  $\eta$  under  $\mathcal{M}$  is expressed using Bayes theorem,

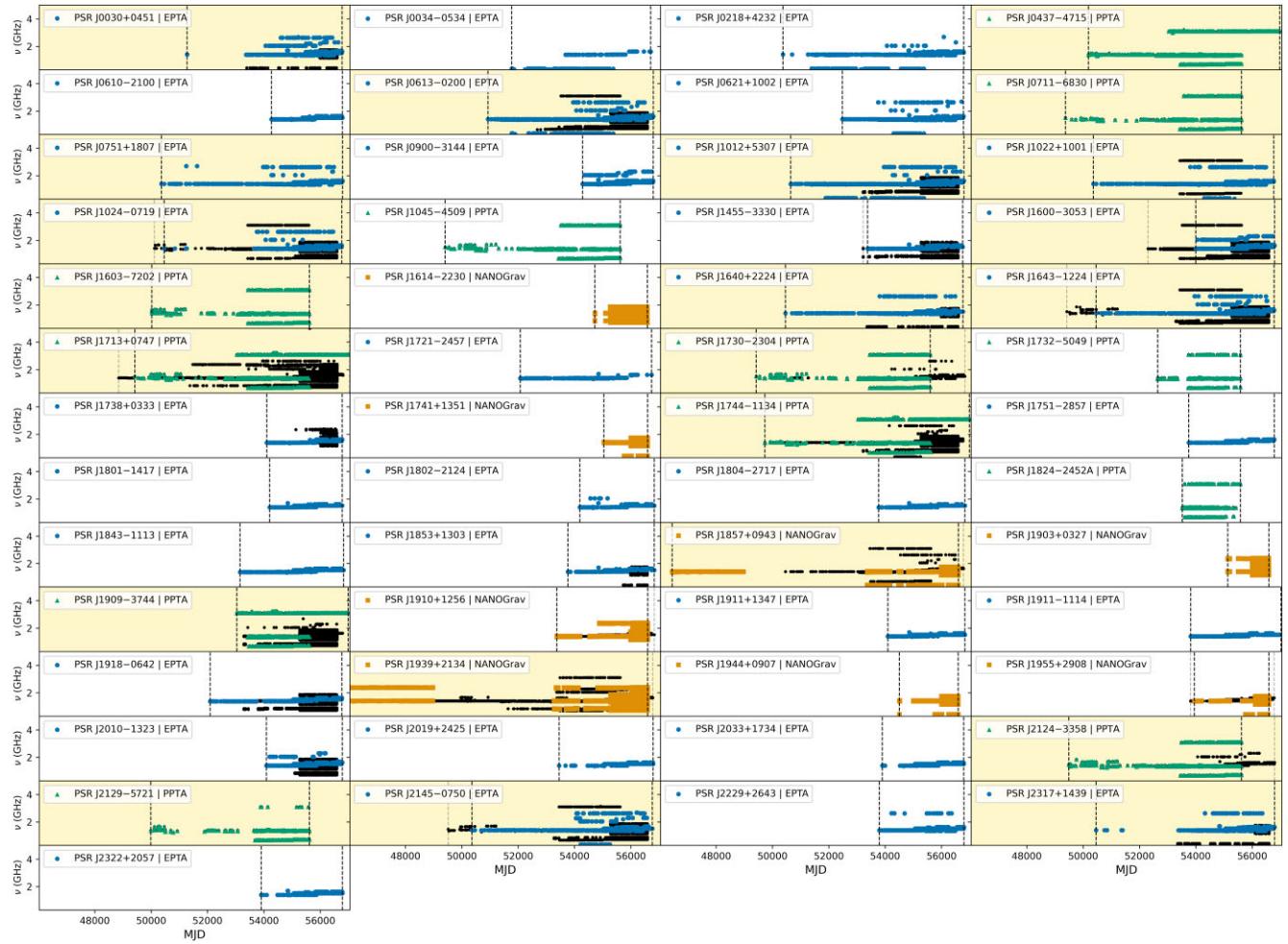
$$\mathcal{P}(\eta | \delta t, \mathcal{M}) = \frac{\mathcal{L}(\delta t | \eta, \mathcal{M})\pi(\eta | \mathcal{M})}{\mathcal{Z}(\delta t | \mathcal{M})}, \quad (3)$$

where  $\pi(\eta | \mathcal{M})$  is our prior probability over  $\eta$  and  $\mathcal{Z}(\delta t | \mathcal{M})$  is the model evidence (or marginal likelihood). Equation (3) is evaluated numerically using Markov chain Monte Carlo (MCMC) or nested sampling. Given two different models  $\mathcal{M}_1$  and  $\mathcal{M}_0$ , the ratio of model evidences is the Bayes factor,

$$\mathcal{B}_{\mathcal{M}_0}^{\mathcal{M}_1} = \frac{\mathcal{Z}(\delta t | \mathcal{M}_1)}{\mathcal{Z}(\delta t | \mathcal{M}_0)}, \quad (4)$$

interpreted as the probability ratio for the data  $\delta t$  under  $\mathcal{M}_1$  versus  $\mathcal{M}_0$ , or equivalently an odds ratio for  $\mathcal{M}_1$  versus  $\mathcal{M}_0$  given the data  $\delta t$  (assuming equal prior odds for both models).  $\mathcal{B}_{\mathcal{M}_0}^{\mathcal{M}_1}$  may be used as a detection statistic for a signal represented by  $\mathcal{M}_1$  if  $\mathcal{M}_0$  is the signal's null hypothesis. For nested models, equation (4) is easily approximated using the Savage-Dickey density ratio (Dickey 1971).

We construct the likelihood and priors using `enterprise` (Ellis et al. 2020) and `enterprise_extensions` (Taylor et al. 2021). We perform parameter estimation using `PTMCMCSampler` (MCMC with parallel tempering; Ellis & van Haasteren 2017) as well as `nutilus` (nested sampling; Lange 2023). We next describe the models used to construct the likelihood and their parameters.



**Figure 2.** All TOAs in DR2 Lite (coloured), compared against all TOAs in Full DR2 (black; Perera et al. 2019), plotted in terms of observation frequency in GHz versus time in MJD for each pulsar. The DR2 Lite TOAs from different PTAs are distinguished as follows: EPTA (blue circles), NANOGrav (orange squares), PPTA (green triangles). Vertical lines show the start and end times of each pulsar’s Lite data set (dashed black), and each pulsar’s combined data set (dashed grey). The 22 pulsar data sets used in EDR2 are highlighted with a gold background. IPTA DR2 contains an additional 44 pulsars not shown, 32 of which with  $> 3$  yr of data were used in the Antoniadis et al. (2022) GWB search.

### 3.1 Timing model

We start with the best-fitting timing model for each pulsar from Perera et al. (2019). The timing model accounts for deterministic delays to a given pulsar’s TOAs accounting for effects such as pulsar spindown, astrometry, binary orbits, dispersion, frequency-dependent pulse profile evolution, and instrumental offsets. However, the presence of time-correlated noise will introduce perturbations to the best-fitting values of the timing model parameters. As such, each pulsar’s timing model is varied using an approximate linearized timing model design matrix  $\mathbf{M} \in \mathcal{T}$ , with elements defined

$$M_{ij} = -\left. \frac{\partial t_i}{\partial \beta_j} \right|_{\beta_{0,j}}, \quad (5)$$

where  $t_i$  is the  $i$ th TOA,  $\beta_j$  is the  $j$ th timing model parameter, and  $\beta_{0,j}$  is the best-fit value of the  $j$ th timing model parameter (van Haasteren & Levin 2013; Taylor 2021). The timing model coefficients  $\epsilon = \beta - \beta_0 \in \mathcal{W}$  are then assigned improper uniform priors, which are implemented numerically as Gaussian priors with (near-)infinite variance, and then marginalized over when computing the likelihood following Johnson et al. (2024).

### 3.2 Noise models

Pulsar noise models account at minimum for white noise, low-frequency red noise, and low-frequency chromatic noise. Here, we update the pulsar noise models for IPTA DR2 from those used in Antoniadis et al. (2022) to reflect recent advances in pulsar noise modelling. In particular, Falxa et al. (2023) recently performed a search for continuous GWs from individual SMBHBs in IPTA DR2 and found that the detailed noise models from Chalumeau et al. (2022), which account for higher frequency sources of noise, were required to mitigate a spurious detection of a continuous GW.

An optimal treatment of pulsar noise would necessitate the creation of fully customized pulsar noise models tailored to IPTA DR2 (e.g. Lentati et al. 2016), but this is beyond the scope of this work. Instead, we use effective pulsar noise models informed by published analyses from individual PTA data sets (Goncharov et al. 2021a; Chalumeau et al. 2022; EPTA Collaboration 2023b; Reardon et al. 2023b; Agazie et al. 2023c; Larsen et al. 2024). We always use  $\log_{10}$ -uniform priors on the amplitude parameters of each noise process. These have been shown to be equivalent to spike and slab priors which enable noise model averaging (van Haasteren 2025). For each

version of the data set, we perform one round of noise analysis in each pulsar for parameter estimation and model validation prior to full-PTA analyses, using the same noise models for each data set. Furthermore, it is also possible to improve these priors by using hierarchical modelling to create a ‘population prior’, which represents the ensemble noise properties of millisecond pulsars (van Haasteren 2024; Goncharov & Sardana 2025). In particular, hierarchical priors have been shown (Goncharov et al. 2024; van Haasteren 2024) to reduce bias in GWB parameter estimation under a scenario where pulsars with similar intrinsic red noise properties become misattributed to the autocorrelations of the GWB (see Goncharov et al. 2022; Zic et al. 2022 for discussions). While this scenario could be relevant for the analysis of IPTA DR2, we do not consider a GWB analysis using hierarchical priors, as the ensemble noise properties obtained from two data sets (i.e. DR2 Lite and Full DR2) will not be equivalent, and comparing results obtained under two different data sets with different priors is not straightforward (best left for future work). None the less, it is useful to compare the ensemble noise properties obtained under different data sets with hierarchical modelling. These results are isolated to Section 4.4, while the remainder of this work uses the standard uninformative priors.

### 3.2.1 Achromatic red noise

Each pulsar’s red noise is modelled as a rank-reduced Gaussian process using a  $N_{\text{TOA}} \times 2N_f$  sine-cosine Fourier design matrix  $\mathbf{F} \in \mathcal{T}$  with elements

$$F_{ij} = \begin{cases} \cos(2\pi f_{j/2} t_i) & \text{for even } j, \\ \sin(2\pi f_{(j-1)/2} t_i) & \text{for odd } j, \end{cases} \quad (6)$$

where we use a linearly spaced frequency basis  $f$ , and  $N_f$  is the number of frequencies used (Lentati et al. 2013). We place a power-law prior on the variance of the Fourier coefficients  $\mathbf{a} \in \mathcal{W}$  at each frequency, parametrized in terms of the power spectral density

$$S(f) = \frac{A_{\text{RN}}^2}{12\pi^2} \left( \frac{f}{\text{yr}^{-1}} \right)^{-\gamma_{\text{RN}}} \text{yr}^3, \quad (7)$$

with uniform priors on the  $\log_{10}$  spectral amplitude at  $f = 1/\text{yr}$   $\log_{10} A_{\text{RN}} \sim \mathcal{U}(-20, -11)$  and spectral index  $\gamma_{\text{RN}} \sim \mathcal{U}(0, 7)$ , while the Fourier coefficients  $\mathbf{a}$  are marginalized over (van Haasteren & Levin 2013). Red noise processes may be reconstructed in the time domain as  $\delta t_{\text{RN}} = \mathbf{F}\mathbf{a}$  by repeated draws from the posterior distribution over  $\mathbf{a}$  (Meyers et al. 2023). During a full-PTA analysis, the frequency basis for intrinsic pulsar red noise is defined to be equivalent to the CRN basis,  $f = (1/T_{\text{DR2}}, \dots, 30/T_{\text{DR2}})$ , where  $T_{\text{DR2}} \cong 30$  yr is the timespan of IPTA DR2, and the number of frequencies are spaced in integer steps of  $1/T_{\text{DR2}}$ . During single pulsar noise analyses, the frequency basis is tailored to the pulsar’s timespan,  $T_{\text{obs}}$ , such that  $f = (1/T_{\text{obs}}, \dots, 30/T_{\text{DR2}})$ . This is chosen because any noise below  $1/T_{\text{obs}}$  in a given pulsar will be degenerate with pulsar spindown parameters. Meanwhile, the truncation frequency  $30/T_{\text{DR2}}$  is chosen to make sure each pulsar’s white noise properties (which could depend on the cutoff if the spectrum is shallow) are consistent across both phases of the analysis. This red noise model is left consistent across all pulsars. However, PSR J1012 + 5307 also exhibits red noise up to very high Fourier modes (Chalumeau et al. 2022; Falxa et al. 2023; EPTA Collaboration 2023b). As such, we add an additional high-frequency power-law red noise process for PSR J1012 + 5307 with  $f = (1/T_{\text{obs}}, \dots, 150/T_{\text{DR2}})$  during all stages of analysis.

### 3.2.2 Chromatic noise

Any time-correlated noise processes depending on the radio-frequency of the pulsar,  $\nu$ , are chromatic. The primary type of chromatic noise is DM noise, varying as  $\delta t_{\text{DM}} \propto \nu^{-2}$ . Similarly to achromatic red noise, we use a Fourier-basis Gaussian process with a power-law prior to model DM noise using hyperparameters  $\log_{10} A_{\text{DM}} \sim \mathcal{U}(-20, -11)$  and  $\gamma_{\text{DM}} \sim \mathcal{U}(0, 7)$ , with an additional scaling  $(\nu/1400 \text{ MHz})^{-2}$  applied to the Fourier design matrix (equation (6)). Following Falxa et al. (2023), we allow the power law frequencies to extend to higher frequencies than achromatic red noise, here using  $f = (1/T_{\text{obs}}, \dots, 150/T_{\text{DR2}})$ . An additional fit for linear and quadratic variations in  $\text{DM}(t)$  are included in all timing models by default.

The solar wind also induces annual quasi-periodic DM variations which we model separately from the Fourier-basis DM Gaussian process. Assuming a spherically-symmetric,  $r^{-2}$  density profile surrounding the Sun, the DM induced by the solar wind is

$$\text{DM}_{\text{solar}}(t) = 4.85 \times 10^{-6} \left( \frac{n_{\text{Earth}}(t)}{\text{cm}^{-3}} \right) \frac{\pi - \theta(t)}{\sin \theta(t)} \text{pc cm}^{-3}, \quad (8)$$

where  $\theta(t)$  is the angle between the Earth-Sun and Earth-pulsar lines of sight, and  $n_{\text{Earth}}(t)$  is the time-dependent solar wind electron density measured at 1 AU from the Sun (You et al. 2007; Hazboun et al. 2022; Nițu et al. 2024). The mean, time-independent component of the electron density is included as a timing model parameter for every pulsar and marginalized over. We additionally fit for time-dependent density perturbations  $\Delta n_{\text{Earth}} \in \mathcal{W}$  along each pulsar’s line-of-sight as a Gaussian process using the model from Nițu et al. (2024), with  $N_b$  equal to the number of pulsar-Sun conjunctions in the pulsar’s data set, and a separate variance parameter sampled for each pulsar using the prior  $\log_{10} \sigma_{n_{\text{Earth}}} \sim \mathcal{U}(-4, 2)$  electrons  $\text{cm}^{-3}$ . Since  $\theta(t)$  is bounded by the ecliptic latitude (ELAT) of each pulsar, many pulsars with large ELAT will be less sensitive to the solar wind, though there may be exceptions depending on radio-frequency coverage and TOA precision (Susarla et al. 2024). We only include the time-dependent model in pulsars for which  $\text{ELAT} < 35^\circ$  and the model is favoured with Savage-Dickey Bayes factor  $\mathcal{B}_0^{\text{SW}(t)} > 1$  using IPTA DR2.

Pulsars may also experience non-dispersive chromatic noise due to effects such as interstellar scattering or pulse profile variability. Scattering results from pulse propagation through an inhomogeneous refractive medium (Cordes & Rickett 1998; Hemberger & Stinebring 2008). A simple model for time-delays introduced by scattering is that  $\delta t \propto \nu^{-\chi}$  with  $\chi = 4.4$ . However, this makes several assumptions, including that the refractive medium is described by Kolmogorov turbulence, the medium is isolated to a thin screen, the pulse is Gaussian, and the pulse broadening function is exponential (Geiger et al. 2025). Violations of these assumptions can and do result in alternative values for  $\chi$  in millisecond pulsars (Turner et al. 2021), especially once transforming from estimates of the scattering delay to the timing residual (Geiger et al. 2025).

Here we account for some of this excess chromatic noise using the same Fourier-basis Gaussian process model as DM noise, except the radio-frequency scaling of the Fourier basis follows  $(\nu/1400 \text{ MHz})^{-\chi}$ , with  $\chi$  as a fit parameter. We incorporate the uncertainty on  $\chi$  in our priors by using a truncated normal distribution,  $\chi \sim \mathcal{N}(4, 0.5) \times \mathcal{U}(2.5, 10)$ , where the lower-bound at  $\chi = 2.5$  prevents degeneracy with DM noise. We include this model for PSRs J0437-4715, J0613-0200, J1600-3053, J1643-1224, J1713+0747, J1903+0327, J1939 + 2134 based on the likely influence of scattering variations in these pulsars’ timing residuals from prior

work (Alam et al. 2021; Srivastava et al. 2023; Reardon et al. 2023b; Agazie et al. 2023d). We also model a chromatic event in PSR J1713 + 0747 using the following deterministic signal (Lam et al. 2018b),

$$s_d(t_i) = -A_d \Theta(t_i - t_0) \exp\left(-\frac{t_i - t_0}{\tau_d}\right) \left(\frac{v_i}{1400 \text{ MHz}}\right)^{-\chi_d}, \quad (9)$$

with uniform priors  $\log_{10} A_d \sim \mathcal{U}(-10, -2)$  s,  $\log_{10} \tau_d \sim \mathcal{U}(0.7, 2.7)$  d,  $t_0 \sim \mathcal{U}(54742, 54768)$  MJD,  $\chi_d \sim \mathcal{U}(1, 6)$  (Goncharov et al. 2021a; Antoniadis et al. 2022). To improve computational efficiency, all chromatic parameters  $\chi$ ,  $\chi_d$  are varied during single pulsar noise analyses, but held fixed to their maximum a posteriori (MAP) values during subsequent full-PTA analyses.

### 3.2.3 White noise

Many white noise parameters are included in IPTA-combined data sets to account for different systematic errors which may be unique to particular observing systems. We apply the same prescription as Antoniadis et al. (2022) for fitting white noise. Two parameter types are diagonal in the  $\mathbf{N}$  matrix: EFAC, which applies a net scaling to the estimated TOA uncertainties, and EQUAD, which adds an additional net uncertainty in quadrature. We also apply ECORR parameters to NANOGrav TOAs, which are intended to model pulse jitter in sub-banded TOAs measured during the same observation epoch using uniform blocks along the diagonal band of the  $\mathbf{N}$  matrix. Separate white noise parameters are applied to TOAs from different receiver and backend combinations in each pulsar, where these combinations are specified by each TOA's -group flag. All white noise parameters are varied during single pulsar noise analyses, and then held fixed to their MAP values during subsequent full-PTA analyses.

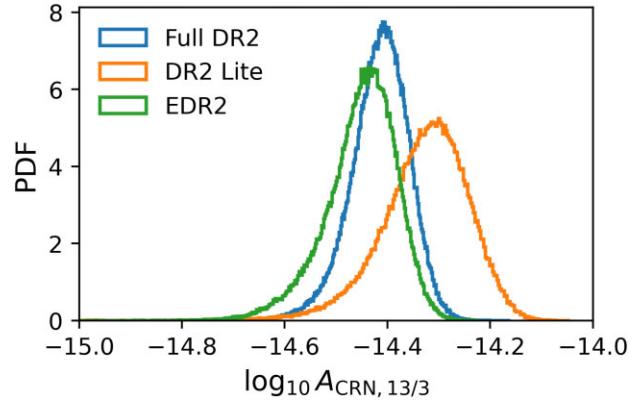
### 3.3 Common signals

In full-PTA analyses, we search for an additional common red noise (CRN) process on top of all components in each pulsar's noise model. The CRN is modelled with the same power-law spectral density as the individual red noise models, equation (7), using new parameters  $A_{\text{CRN}}$  and  $\gamma_{\text{CRN}}$  which are fit for simultaneously in all pulsars at once. In a single pulsar analysis, the achromatic red noise includes contributions from both common and intrinsic pulsar noise. Switching from the single pulsar to full-PTA analysis decouples the total achromatic red noise into the separate intrinsic and common channels. The CRN also uses the same frequency basis  $f$  for all pulsars. To be consistent with Antoniadis et al. (2022), we use  $N_{\text{freqs}} = 13$  components for a frequency grid  $f = (1/T_{\text{DR2}}, \dots, 13/T_{\text{DR2}})$  for each analysis.

To model cross-correlations in a full-PTA analysis we define the cross-power spectral density,

$$S_{ab}(f) = \Gamma_{ab} S(f), \quad (10)$$

where  $\Gamma_{ab}$  is the overlap reduction function (ORF) encoding the geometric cross-correlation between pulsars  $a$  and  $b$  as a function of their sky-separation angle, and  $S(f)$  is the power spectral density in each pulsar, given by the form of equation (7) if assuming a power law spectrum. One can specify different signals by the form of the ORF:  $\Gamma_{ab} = \delta_{ab}$  represents a purely autocorrelated CRN process (uncorrelated between pulsars), whereas  $\Gamma_{ab}$  given by the Hellings-Downs curve is the signature of an isotropic GWB under general relativity. Alternative ORFs given by monopolar and dipolar forms in pulsar sky separation angle would result from errors in terrestrial



**Figure 3.** Comparison of the posterior PDF for  $\log_{10} A_{\text{CRN}}$  at fixed  $\gamma_{\text{CRN}} = 13/3$  from Full DR2 (blue, middle distribution), DR2 Lite (orange, rightmost distribution), and EDR2 (green, leftmost distribution). While the posteriors from all three data sets show evidence of a CRN, the amplitude measured using DR2 Lite is shifted to larger values than the posteriors from the combined data sets.

time standards (Hobbs et al. 2012) and conversion to the Solar system barycenter (Champion et al. 2010), respectively.

## 4 IPTA DR2 LITE VERSUS COMBINED ANALYSIS RESULTS

Here we present the results of our common signal search and analysis of DR2 Lite (Table 1, row 2), in comparison with the same analysis of Full DR2, which was originally carried out in Antoniadis et al. 2022 (Table 1, row 1). We also perform an analysis on the fully-combined data set with only 22 pulsars, designated here as EDR2, that could reflect an intermediate stage of the data combination process (Table 1, row 3).

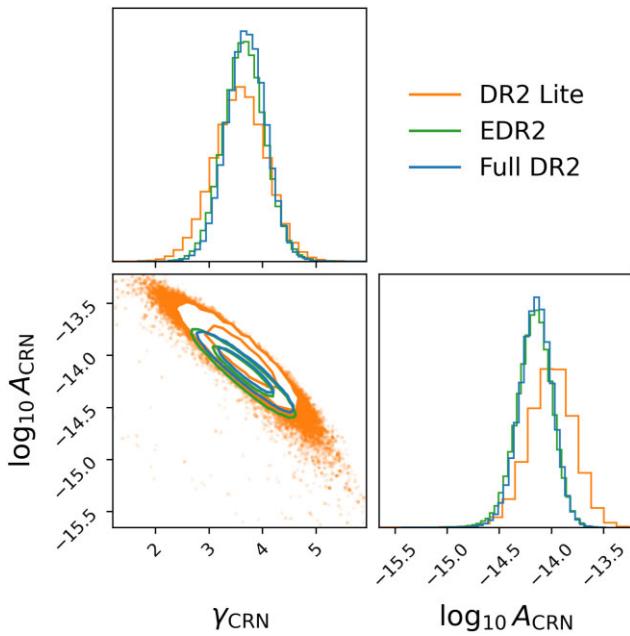
### 4.1 Common red noise

We first compare our inferences on the CRN parameters using each data set to see how much information we can learn about the common signal using the Lite method, and how much our inferences improve using the combined data.

#### 4.1.1 Full-PTA parameter estimation

First we perform a simultaneous analysis of all pulsars in each data set to estimate the CRN parameters assuming a fixed spectral index  $\gamma_{\text{CRN}} = 13/3$ . Fig. 3 compares the posterior PDFs for  $\log_{10} A_{\text{CRN}, 13/3}$  from each data set. We also perform the same comparison using a varied  $\gamma_{\text{CRN}}$  model, for which we show the CRN posteriors in Fig. 4. The median and 95 per cent credible intervals on CRN parameters from both models are reported in Table 2. We further report in Table 2 the upper limit  $A_{\text{CRN}, 13/3}^{95 \text{ per cent}}$  for each data set, estimated as the 95 per cent one-sided Bayesian credible interval after replacing the  $\log_{10}$ -uniform priors on  $A_{\text{CRN}, 13/3}$  with uniform priors.

We find that all three data sets, including DR2 Lite, are able to detect a CRN, however, there are differences in spectral characterization. In particular, the amplitude measured using DR2 Lite is very large—systematically higher than amplitude measured from the combined data sets, with a median amplitude  $A_{\text{CRN}, 13/3}$  measured 23 per cent larger using DR2 Lite than it is using Full DR2. This implies that DR2 Lite is allowing excess noise intrinsic to the pulsars



**Figure 4.** Comparison of posteriors for the CRN parameters  $\log_{10} A_{\text{CRN}}$  and  $\gamma_{\text{CRN}}$  from Full DR2 (blue, narrowest distribution), DR2 Lite (orange, widest distribution), and EDR2 (green, aligned with Full DR2). Adding more data in the combination results in tighter constraints on the CRN parameters. Contours enclose 68 per cent and 95 per cent of the posterior from each data set (with samples outside the 95 per cent credible region shown only for DR2 Lite).

to leak into the common channel (Zic et al. 2022). The upside then is that performing data combination apparently helps to mitigate this effect. We explore this discrepancy further in Section 4.1.4.

Additionally, Fig. 4 shows that the CRN parameter constraints become more precise as more data are added. To quantify the improvement in precision, we estimate the area  $\mathcal{A}$  of the 2D region in  $\log_{10} A_{\text{CRN}}, \gamma_{\text{CRN}}$  enclosed within 95 per cent of the posterior from each data set. We find the ratio of areas with respect to Full DR2 to be  $\mathcal{A}_{\text{Lite}}/\mathcal{A}_{\text{Full}} = 2.25$  and  $\mathcal{A}_{\text{EDR2}}/\mathcal{A}_{\text{Full}} = 1.23$ , i.e. Full DR2 is 2.25 times *more* precise at spectral characterization than DR2 Lite, and 1.23 times more precise than EDR2.

Fig. 5 further compares each data set using a more generic ‘free-spectral’ model for the CRN, where we drop the assumption of a power-law spectrum and sample the timing residual power at each discrete frequency  $f_i$  as independent parameters with prior  $\log_{10} \rho_i \sim \mathcal{U}(-10, -4)$  in units of seconds (Lentati et al. 2013). The top panel shows the posteriors on each  $\log_{10} \rho_i$  from each data set, which plotted versus frequency represent the amplitude spectrum of

the CRN. Here, it is valid to compare amplitude spectra from the different data sets as they share the same baseline  $T_{\text{DR2}}$ , otherwise the amplitude spectral density would be the relevant quantity to compare. Overplotted lines depict median *power law* spectra from each data set, for comparison. The bottom panel of Fig. 5 shows the  $\log_{10}$  Bayes factors,  $\log_{10} \mathcal{B}_{\text{CRN}-\rho_i}^{\text{CRN}}$ , for the free spectral CRN model versus the same model without the inclusion of common power in the frequency bin at frequency  $f_i$ , as measured using the Savage-Dickey density ratio. These quantify the detection significance of the CRN at each frequency bin, where  $\log_{10} \mathcal{B}_{\text{CRN}-\rho_i}^{\text{CRN}} > 0$  indicates the data do favour the inclusion of additional common power at frequency  $f_i$ .

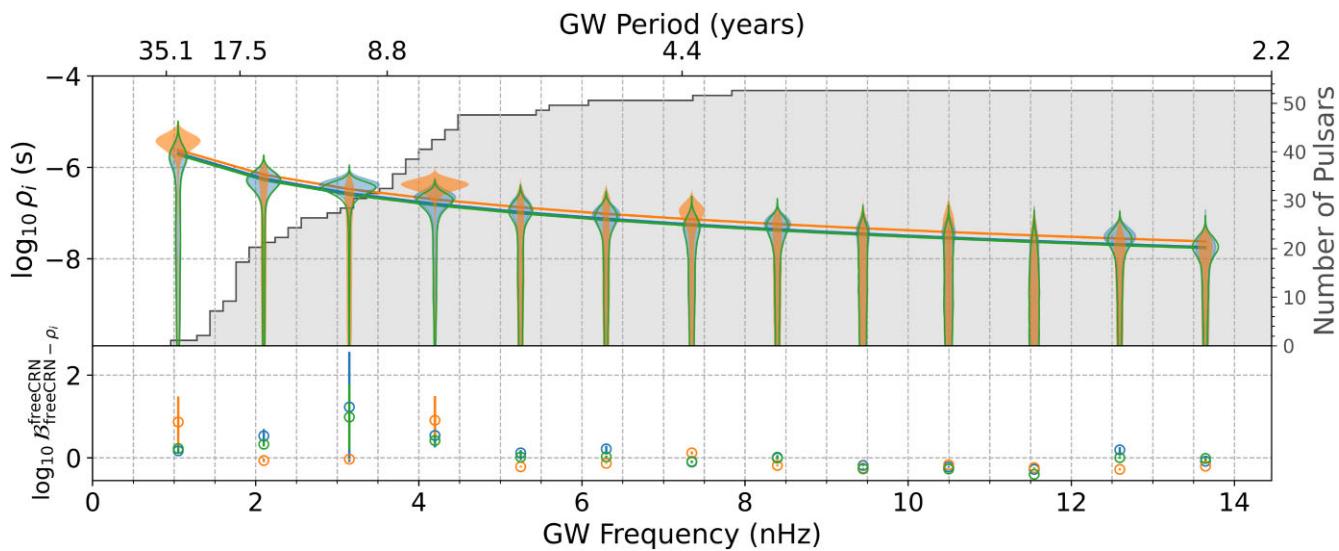
EDR2’s amplitude spectrum is comparable to Full DR2, with small deviations in the posteriors at higher frequencies. Both spectra are consistent with a power law form and display strongest detections of power in the 2 – 3 nHz range. Meanwhile, the posterior power in DR2 Lite’s amplitude spectrum exceeds that from the combined data sets at several frequencies, which is consistent with the higher power-law amplitude measured using DR2 Lite. We focus on the 1 and 4 nHz frequencies (corresponding to periods of  $\sim 30$  and  $\sim 7.5$  yr, respectively) where power is detected *more significantly* using DR2 Lite than the combined data sets. Relatively few pulsars in IPTA DR2 have sufficiently long timespans to contribute meaningfully to constraining the posterior at 1 nHz. The Fig. 5 cumulative histogram of pulsars over their observation timespan  $1/T_{\text{obs}}$  illustrates this effect: Only 3 pulsars (PSRs J1713+0747, J1857+0943, J1939+2134) have  $T_{\text{obs}} > 21$  yr, 2 of which (PSRs J1857+0943 and J1939 + 2134) have very large data gaps spanning  $\sim 10$  yr in DR2 Lite, as shown in Fig. 2. Unconstrained noise in even 1 of these pulsars using DR2 Lite could feasibly cause changes in timing residual power at 1 nHz. As for the 4 nHz posterior, it visibly deviates from the median power law fit obtained from Full DR2 and EDR2, likely producing the large amplitude measured in the power law analysis. While 41 pulsars have sufficient timespan to resolve this frequency, details in the noise characterization of a minority of especially sensitive pulsars may still produce such features in the common spectrum (Hazboun et al. 2020; Larsen et al. 2024).

#### 4.1.2 Factorized likelihood analysis

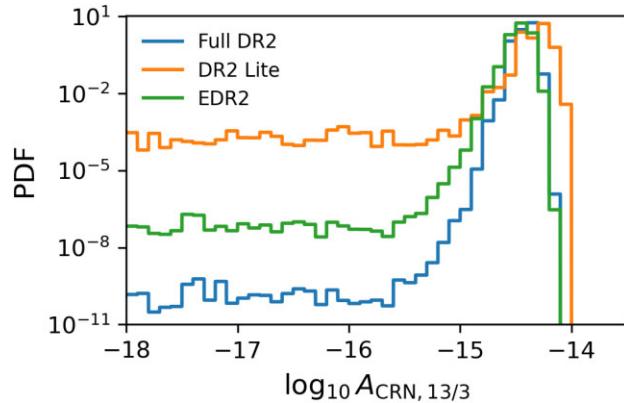
We gain additional information about the significance of an autocorrelated common process using the Factorized Likelihood (Taylor et al. 2022). Using the factorized likelihood, the likelihood and common noise posteriors are estimated as a product of these statistics as measured from analyses of individual pulsars. The factorized likelihood not only enables rapid PTA analyses but also allows the analysis of different Lite data sets at no additional computational cost, as demonstrated in Agazie et al. (2024). Using the factorized likelihood, we assign the CRN model a fixed spectral index  $\gamma_{\text{CRN}} = 13/3$

**Table 2.** CRN parameters and statistics as a function of data set. Using the fixed  $\gamma_{\text{CRN}} = 13/3$  model, we report median and 95 per cent credible intervals on  $A_{\text{CRN},13/3}$  (using  $\log_{10}$ -uniform priors), 95 per cent upper limits  $A_{\text{CRN},13/3}^{95 \text{ per cent}}$  (using uniform priors), and Bayes factors on the CRN  $\mathcal{B}_{\text{CRN},13/3}^{95 \text{ per cent}}$  estimated using the Savage-Dickey density ratio from a factorized likelihood analysis. Using the varied  $\gamma_{\text{CRN}}$  model, we report median and 95 per cent credible intervals on  $A_{\text{CRN}}$  and  $\gamma_{\text{CRN}}$ . As data are progressively added from Lite to full, CRN detection statistics become more significant and parameter estimates become more precise, while the upper limit decreases with combined data.

IPTA DR2 subset	Fixed $\gamma_{\text{CRN}} = 13/3$		Varied $\gamma_{\text{CRN}}$		
	$A_{\text{CRN},13/3}$	$A_{\text{CRN},13/3}^{95 \text{ per cent}}$	$\log_{10} \mathcal{B}_{\text{CRN},13/3}^{95 \text{ per cent}}$	$A_{\text{CRN}}$	$\gamma_{\text{CRN}}$
DR2 Lite	$4.8_{-1.8}^{+1.8} \times 10^{-15}$	$6.4 \times 10^{-15}$	$3.0 \pm 0.2$	$10.0_{-6.5}^{+15.6} \times 10^{-15}$	$3.6_{-1.0}^{+0.9}$
EDR2	$3.6_{-1.2}^{+1.0} \times 10^{-15}$	$4.5 \times 10^{-15}$	$6.4 \pm 0.2$	$6.9_{-3.9}^{+7.3} \times 10^{-15}$	$3.7_{-0.8}^{+0.8}$
Full DR2	$3.9_{-1.0}^{+1.0} \times 10^{-15}$	$4.7 \times 10^{-15}$	$9.1 \pm 0.3$	$7.2_{-3.8}^{+7.0} \times 10^{-15}$	$3.7_{-0.7}^{+0.7}$



**Figure 5.** Comparison of the inferred CRN amplitude spectrum,  $\rho(f) = \sqrt{S(f)/T_{\text{DR2}}}$ , using Full DR2 (blue, outlined, shaded), DR2 Lite (orange, shaded, no outlines), and EDR2 (green, outlined, hollow). The violins in the top panel display posteriors on the  $\log_{10}$  amplitude of common timing residual power  $\rho_i$  at each discrete frequency  $f_i$ , and the solid lines indicate power law spectra corresponding to the median values of the CRN parameters from Fig. 4. To understand the number of pulsars contributing at each frequency, the violins are shown over a cumulative histogram of the pulsars in IPTA DR2 based on their observation timespans  $f = 1/T_{\text{obs}}$ . The bottom panel quantifies the detection significance of common power at each bin via the  $\log_{10}$  Bayes factor for a model containing the full CRN amplitude spectrum versus the same model with a single free spectral bin removed at the corresponding frequency. Using DR2 Lite,  $\log_{10} \rho_i$  is measured most significantly at  $f \sim 1$  and  $f \sim 4$  nHz.



**Figure 6.** Inferences of the CRN amplitude at  $\gamma = 13/3$  using the factorized likelihood with each data set: Full DR2 (blue, smallest tail), DR2 Lite (orange, largest tail), EDR2 (green). The factorized likelihood allows inference of the low-amplitude of each distribution, which corresponds to a null-detection. The probability of a null-detection is highest with DR2 Lite and lowest with Full DR2, as expected when using more data in the analysis. EDR2 achieves greater detection significance than DR2 Lite despite including fewer pulsars.

in each pulsar, while the intrinsic red noise model's spectral index is allowed to vary (otherwise, the intrinsic red noise and CRN signals are completely degenerate). Since the CRN may always be assigned to the intrinsic RN model during each single-pulsar analysis, the CRN amplitude in each pulsar is always finite at  $A_{\text{CRN},13/3} < 10^{-17}$ , therefore the Savage-Dickey density ratio can always be used to estimate the CRN Bayes factor from the product.

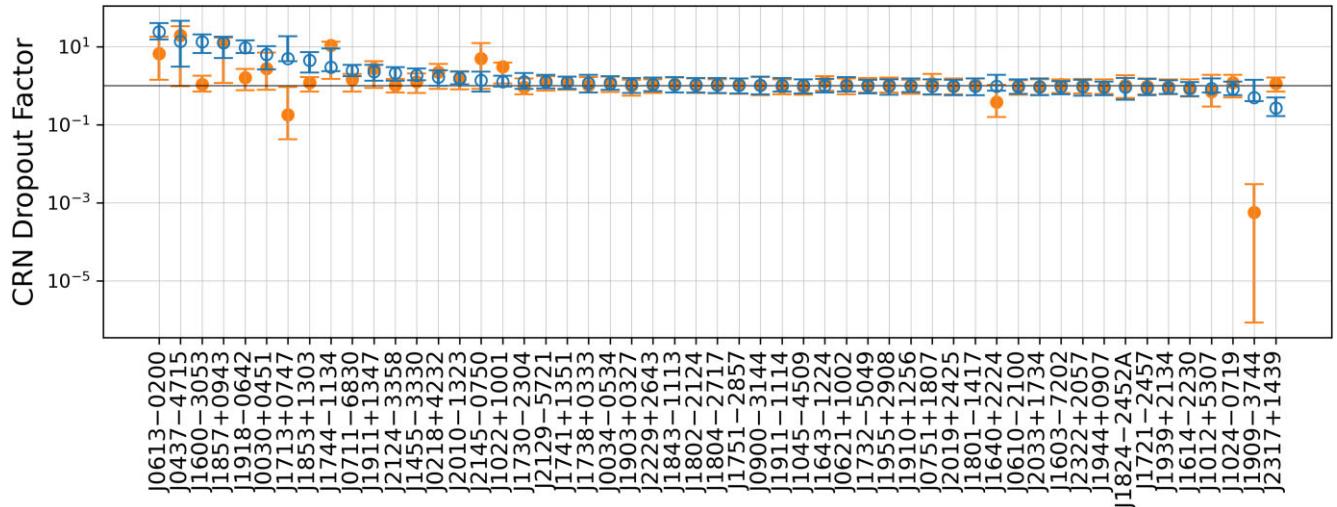
Fig. 6 displays the posteriors over  $\log_{10} A_{\text{CRN},13/3}$  obtained from the factorized likelihood analysis, and Column 5 of Table 2 reports the Bayes factors for a model containing a CRN with  $\gamma_{\text{CRN}} = 13/3$  versus a model with no CRN. We estimate the Bayes factor means and uncertainties based on the distributions of probabilities within

$\log_{10} A_{\text{CRN},13/3} < -16$  shown in Fig. 6. All three data sets' Bayes factors are of  $\mathcal{O}(10^3)$  or greater, signifying strong preference for a CRN, in agreement with the full-PTA analysis. However, the results show progressively more support for a CRN as the data combination process continues, with the odds of a CRN increasing by a factor of  $\sim 10^{3.4}$  going from DR2 Lite to EDR2, and again by a factor of  $\sim 10^{2.7}$  going from EDR2 to Full DR2. This result is unsurprising as we add more data into the analysis, but none the less underscores the effectiveness of data combination for improving our detections of common signals.

#### 4.1.3 Dropout analysis

The factorized likelihood can be also used to rapidly compute dropout factors (DFs), which are Bayes factors for a model with the CRN in  $N$  pulsars versus a model with the CRN in  $N - 1$  pulsars, while the excluded pulsar  $p$  is modelled with intrinsic noise only (Arzoumanian et al. 2020b; Antoniadis et al. 2022; Taylor et al. 2022). If pulsar  $p$ 's DF  $> 1$ , then pulsar  $p$ 's data support the CRN measurement (i.e. the CRN Bayes factor would decrease by the DF upon removing pulsar  $p$  from the array), whereas if pulsar  $p$ 's DF  $< 1$  then pulsar  $p$ 's data are in tension with the CRN (i.e. the CRN Bayes factor would increase by the DF upon removing pulsar  $p$  from the array). Comparing the DFs across pulsars is a useful way to diagnose which pulsars most heavily favour or disfavour the inclusion of a CRN. To see how individual pulsars affect parameter estimation on  $A_{\text{CRN},13/3}$ , see Section 4.1.4.

Fig. 7 compares each pulsar's DF from DR2 Lite and Full DR2, as computed using the factorized likelihood. We follow the procedure from Antoniadis et al. (2022) to compute uncertainties on the DFs: first we vary the analysis over 25 combinations of metaparameters and then for each combination, we generate 1000 bootstrap realizations by re-sampling the chain with replacement. The Full DR2 DFs are consistent with those reported in Antoniadis et al. (2022), with



**Figure 7.** Comparison of CRN dropout factors (DFs) computed using the factorized likelihood for each pulsar, as computed using DR2 Lite (orange) and Full DR2 (blue). DFs help diagnose which pulsars are most responsible for the results of a CRN search, where pulsars with  $DF > 1$  support the presence of a CRN and pulsars with  $DF < 1$  are in tension with a CRN. Pulsars are ranked from highest to lowest DF using Full DR2, and errors are estimated as the 95 per cent confidence interval of 25 000 bootstrap realizations, following Antoniadis et al. (2022). PSR J1909–3744’s especially low DF signifies its total measured noise has a lower amplitude than the CRN, as further shown in Fig. 11.

several pulsars’ DFs  $> 1$ . Using DR2 Lite, fewer DFs are found to be larger than 1, which is unsurprising as the Bayes factor for the CRN is lower. PSR J1909–3744’s value of  $DF = 6^{+24}_{-6} \times 10^{-4}$  using DR2 Lite is extremely small (even within bootstrapping errors), as such this pulsar seems to be in major tension with the CRN process when using DR2 Lite. This is highly unusual, but may be explained if PSR J1909–3744 has a much lower upper limit on the CRN than the remaining pulsars, as discussed in Section 4.1.4. Using Full DR2, PSR J1909–3744’s DF is  $DF = 0.49^{+0.91}_{-0.08}$ , indicating PSR J1909–3744 is now more in line with the CRN and with the other combined pulsars. We follow up further with PSR J1909–3744 in Section 4.3. PSR J1713+0747 is also an interesting case, switching from disfavouring to supporting the CRN, going from  $DF = 0.18^{+0.73}_{-0.14}$  using DR2 Lite to  $DF = 4.93^{+13.34}_{-0.72}$  using Full DR2. Overall, five pulsars experience significant (i.e.  $> 2\sigma$ ) boosts to their DF after data combination, while only PSR J2317+1439’s DF is significantly higher using DR2 Lite.

#### 4.1.4 Assessing the large common amplitude in DR2 Lite

While the dropout analysis in Section 4.1.3 shows that the *detection significance* of the CRN may be skewed by individual outlier pulsars using DR2 Lite, it still leaves open the question why the CRN amplitude is systematically larger when using DR2 Lite as opposed to Full DR2. To investigate further, we apply a modified version of the dropout analysis which compares how the *discrepancy in the measured amplitude* depends on which pulsars are dropped. To quantify this discrepancy, we compute a distribution over the shift or difference between the values for  $A_{\text{CRN},13/3}$  from each data set,  $\Delta A = A_{\text{Lite}} - A_{\text{Full}}$ , the mean of which is defined

$$\langle \Delta A \rangle = \int dA_{\text{Lite}} \int dA_{\text{Full}} (A_{\text{Lite}} - A_{\text{Full}}) \times \mathcal{P}(A_{\text{Lite}} | \delta t_{\text{Lite}}) \mathcal{P}(A_{\text{Full}} | \delta t_{\text{Full}}) \quad (11)$$

$$\approx \frac{1}{N_{\text{samples}}} \sum_{i,j} (A_{\text{Lite}})_i - (A_{\text{Full}})_j. \quad (12)$$

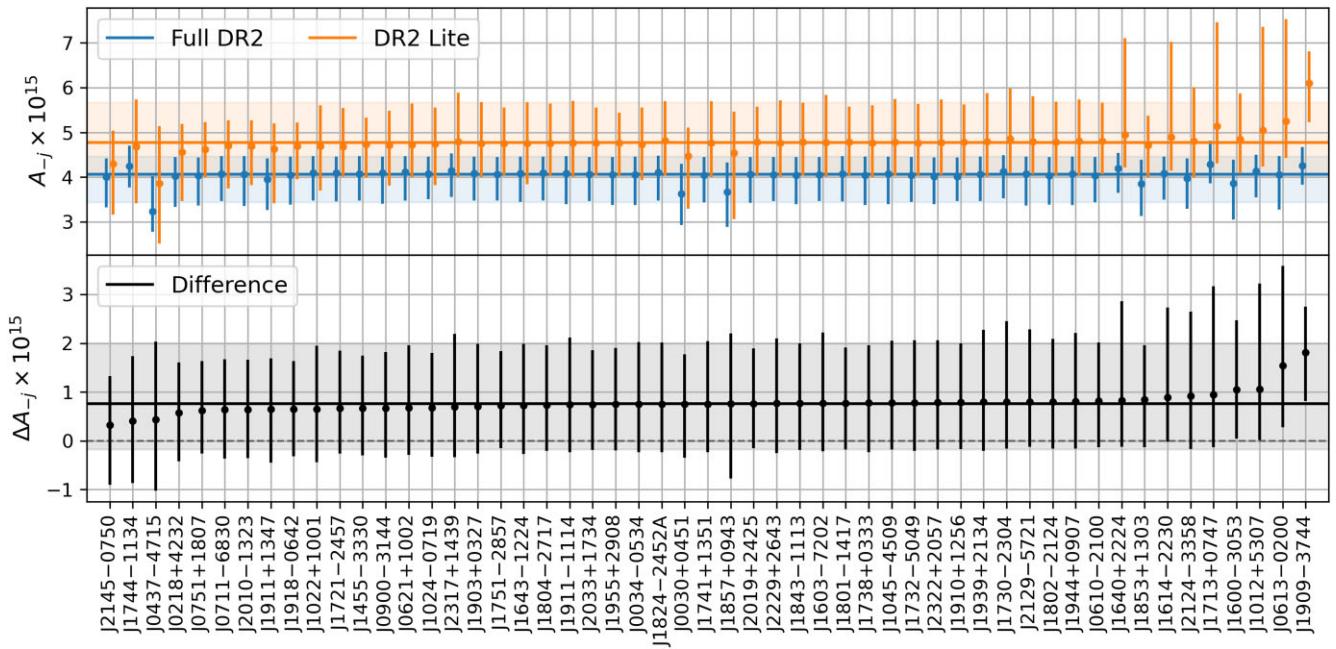
For brevity we have omitted the previously used subscripts so that  $A \equiv A_{\text{CRN},13/3}$ , and line 12 makes explicit that we compute this difference simply by drawing Monte Carlo pairs of samples from the original factorized likelihood posteriors from each data set. This quantity is related to the tension metric from Agazie et al. (2024). When including all pulsars, the median of this difference comes out to  $\Delta A = 0.8^{+1.2}_{-1.0} \times 10^{-15}$ , where the quantiles enclose 68 per cent of the distribution about the median. The value  $\Delta A = 0$  is only just within the 68 per cent credible region, indicating a marginal tension on the level of  $0.81\sigma$ .

To test the dependence of this discrepancy on individual pulsar data sets, we recompute the CRN amplitude at  $\gamma_{\text{CRN}} = 13/3$  via the factorized likelihood method (equation 5 from Taylor et al. 2022) but with pulsar  $j$  removed from the data set, i.e.

$$p_{-j}(\log_{10} A | \{\delta t\}_{-j}) \equiv p \left( \log_{10} A | \prod_{i \neq j}^{N_{\text{psr}}} \delta t_i \right) \propto \prod_{i \neq j}^{N_{\text{psr}}} p \left( (\log_{10} A)_i | \delta t_i \right), \quad (13)$$

where  $i$  indexes each pulsar’s data set  $\delta t_i$ , and the  $-j$  subscripts are a shorthand to denote pulsar  $j$  has been dropped. We compute  $p_{-j}(\log_{10} A | \{\delta t\}_{-j})$  for every pulsar  $j$  using both DR2 Lite and Full DR2, at which points combining with equation (12) yields a new measure of the discrepancy,  $\Delta A_{-j}$ , with the  $j$ th pulsar removed from both data sets.

The bottom panel of Fig. 8 shows our estimates of  $\Delta A_{-j}$  for each pulsar  $j$  individually removed by themselves. Shaded regions show the corresponding estimates  $\Delta A$  with no pulsars removed. To better understand what is causing any measured differences in  $\Delta A_{-j}$ , we also show the estimates on  $A_{-j}$  using each data set in the top panel of Fig. 8 (and corresponding shaded regions for  $A$  with no pulsars removed). If  $A_{-j} > A$ , then pulsar  $j$  is driving down the amplitude estimate, whereas if  $A_{-j} < A$ , then pulsar  $j$  is adding power to the common amplitude.



**Figure 8.** Parameter estimation on  $A_{\text{CRN},13/3}$  using the factorized likelihood posteriors with each pulsar  $j$  dropped out one at a time by themselves. The top panel shows medians and 68 per cent credible intervals over  $A_{-j}$  (equation 13) using each data set, while the bottom panel shows the corresponding difference  $\Delta A_{-j}$  between DR2 Lite and Full DR2 (equation 12). For comparison, the solid lines and shaded regions show the corresponding estimates  $A$  and  $\Delta A$  with no pulsars removed. Pulsars are ranked from the lowest to highest median values of  $\Delta A_{-j}$ . Several individual pulsars, when removed by themselves, skew the estimates of  $\Delta A$ , showing the overall discrepancy is sensitive to systematic errors in the individual pulsars.

There are several pulsars, when excluded from the analysis, that reduce the discrepancy in  $\Delta A$ . The most impactful pulsars on this end are PSRs J2145–0750, J1744–1134, and J0437–4715; their individual removal reduces the discrepancy to a  $0.26\sigma$ ,  $0.32\sigma$ , and  $0.30\sigma$  level, respectively. Interestingly, each pulsar skews the distributions  $A_{-j}$  in a different way. PSR J2145–0750 displays very loud red noise with a steep power law index  $\gamma > 4$  in IPTA DR2 (Caballero et al. 2016; Perera et al. 2019); in Full DR2, this is decoupled from the CRN, whereas in DR2 Lite, the red noise is not sufficiently resolved from DM variations (see Table 4; Section 4.3) nor is it resolved from the CRN, causing the red noise to pollute the CRN process. PSR J1744–1134, among the best timers in the data set, affects the discrepancy simply by reducing  $A_{\text{CRN},13/3}$  using Full DR2, whereas there is less constraint on its noise using DR2 Lite (Table 4). Finally, PSR J0437–4715, which only has TOAs from the PPTA in IPTA DR2, strongly drives up the CRN amplitude for both data sets. Notably, PSR J0437–4715 has the highest FoM behind PSR J1939+2134 and therefore has a disproportionately strong effect on the analysis, on top of very challenging noise properties to model due to its brightness, which may well contribute additional systematic error in this analysis (see e.g. Lentati et al. 2016; Goncharov et al. 2021a; Reardon et al. 2023b, 2024). In summary, we consider the discrepancy in the CRN amplitude between DR2 Lite and Full DR2 to be statistically insignificant as it is not robust to outliers among the individual pulsars.

Finally, there are a few pulsars on the opposite end where the discrepancy *widens* with their removal. This is unsurprising if an equal number of pulsars also narrow the discrepancy. However, this analysis sheds insight into the nature of PSR J1909–3744’s extremely low DF from Section 4.1.3. We see on the far right-hand side of Fig. 8 that removing PSR J1909–3744 from the analysis results in a much *larger* CRN amplitude of  $A_{-j} = 6.1^{+1.4}_{-1.5} \times 10^{-15}$

**Table 3.** Detection statistics for Hellings-Downs correlations as a function of the data set. The middle column reports Bayes factors for an Hellings-Downs cross-correlated model versus an autocorrelated CRN model, as estimated using both likelihood reweighting and the HyperModel. The last column reports the Bayesian S/N for Hellings-Downs correlations from a noise marginalized multiple component optimal statistic (OS) analysis.

IPTA DR2 subset	$\mathcal{B}_{\text{CRN},13/3}^{\text{HD},13/3}$	S/N <sub>Bayes</sub>
	Reweighting	HyperModel
DR2 Lite	$0.662 \pm 0.005$	$0.57 \pm 0.03$
EDR2	$2.841 \pm 0.006$	$3.0 \pm 0.2$
Full DR2	$1.39 \pm 0.04$	$1.45 \pm 0.07$
		OS

(with errors enclosing 95 per cent quantiles). This shows PSR J1909–3744’s DF is so low because it is uniquely suppressing a higher-amplitude mode of the CRN posterior, which is more likely to be dominated by intrinsic pulsar noise. Meanwhile, the Full DR2 amplitude  $A_{-j}$  when dropping PSR J1909–3744 is shifted by much less using Full DR2, reflective of the fact that several more pulsars contribute constraints on the CRN measurement after performing data combination.

## 4.2 Cross-correlations

In order for the data to signify evidence of a GWB, they ought to favour a cross-correlated common signal following the Hellings-Downs curve, as opposed to a purely autocorrelated common signal. We assume that data combination will ultimately improve our ability to resolve a cross-correlated GWB over the use of uncombined data. IPTA DR2 has limited value to test this assumption, since the statistics for cross-correlations in IPTA DR2 obtained first by

**Table 4.** Bayes factors for different processes computed from single pulsar analyses. The 24 pulsars featured here are those from IPTA DR2 with data from 2 or more PTAs.  $\mathcal{B}_0^{\text{PLRN}}$  is the Bayes factor for a model with power-law red noise included versus the same model without,  $\mathcal{B}_0^{\text{PLDM}}$  for power-law DM noise,  $\mathcal{B}_0^{\text{PLChr}}$  for higher-order chromatic noise, and  $\mathcal{B}_0^{\text{SWG}}$  for Gaussian perturbations to the solar wind electron density along the Earth-pulsar line of sight. These Bayes factors were computed as the Savage-Dickey density ratio using each signal's relevant amplitude parameter; where there were a lack of samples in the tail of each parameter's posterior, we place a lower bound of  $\mathcal{B} > 10^3$ . Bold values mark cases where  $\mathcal{B} > 10$ , indicating strong evidence for the noise process. The total number of bolded entries are tallied at the bottom of each column. Entries with asterisks (\*) indicate cases of severe covariance within the noise model parameter space, such that either, but not both, marked signals are favoured by the data (i.e. the data cannot distinguish the chromaticity of the process). The daggered ( $\dagger$ ) entries for PSR J1012+5307 indicate that the high-frequency achromatic red noise process was used to compute the Bayes factor, rather than the 30 frequency red noise process.

Pulsar	$\mathcal{B}_0^{\text{PLRN}}$	DR2 Lite single pulsar Bayes factors	$\mathcal{B}_0^{\text{PLDM}}$	$\mathcal{B}_0^{\text{PLChr}}$	$\mathcal{B}_0^{\text{SWG}}$	$\mathcal{B}_0^{\text{PLRN}}$	Full DR2 single pulsar Bayes factors	$\mathcal{B}_0^{\text{PLDM}}$	$\mathcal{B}_0^{\text{PLChr}}$	$\mathcal{B}_0^{\text{SWG}}$
J0030+0451	2.1	1.0	—	—	—	$> 10^3$	3.6	—	—	—
J0613-0200	3.2*	2.0*	1.2*	—	—	$> 10^3$	9.0	<b>19.8</b>	—	—
J1012+5307	1.1 <sup>†</sup>	0.6	—	—	—	$> 10^3\dagger$	$> 10^3$	—	—	—
J1022+1001	1.2*	3.0*	—	0.8	1.4	<b>56.8</b>	—	—	<b>386.0</b>	—
J1024-0719	1.7*	1.3*	—	—	$> 10^3$	$> 10^3$	—	—	—	—
J1455-3330	1.4	1.4	—	—	0.8	0.7	—	—	—	—
J1600-3053	0.8	1.2*	1.6*	1.7	<b>15.7</b>	3.0	$> 10^3$	—	<b>292.7</b>	—
J1640+2224	0.7	0.7	—	—	1.1	$> 10^3$	—	—	—	—
J1643-1224	0.7	2.9*	1.0*	—	$> 10^3$	$> 10^3$	$> 10^3$	—	—	—
J1713+0747	<b>67.6</b>	1.0	0.9	—	$> 10^3$	$> 10^3$	$> 10^3$	$> 10^3$	—	—
J1730-2304	0.8	1.1	—	—	0.8	0.7	—	—	—	—
J1738+0333	0.9	0.8	—	—	1.2*	1.8*	—	—	—	—
J1744-1134	1.5*	1.1*	—	—	$> 10^3$	<b>50.0</b>	—	—	—	—
J1853+1303	0.8	0.8	—	—	1.4	1.1	—	—	—	—
J1857+0943	9.6	$> 10^3$	—	—	<b>1087.1</b>	$> 10^3$	—	—	—	—
J1909-3744	0.5	$> 10^3$	—	—	$> 10^3$	$> 10^3$	—	—	—	—
J1910+1256	0.8	<b>30.9</b>	—	—	0.8	<b>13.7</b>	—	—	—	—
J1918-0642	1.2*	1.9*	—	4.1	2.3	$> 10^3$	—	—	4.0	—
J1939+2134	$> 10^3$	$> 10^3$	$> 10^3$	—	$> 10^3$	$> 10^3$	$> 10^3$	—	—	—
J1955+2908	0.8	<b>24.6</b>	—	—	0.9	<b>169.4</b>	—	—	—	—
J2010-1323	1.5	1.0	—	—	1.4	$> 10^3$	—	—	—	—
J2124-3358	0.9	0.8	—	3.7	0.9	0.9	—	—	1.0	—
J2145-0750	2.2*	3.5*	—	0.8	$> 10^3$	<b>158.0</b>	—	—	7.0	—
J2317+1439	0.7	0.7	—	—	0.6	$> 10^3$	—	—	—	—
Total $\mathcal{B} > 10$ :	2	5	1	0	12	16	5	2		

Antoniadis et al. (2022) are well below the thresholds required for GW detection (Allen et al. 2023). Thus, the following comparisons of cross-correlation statistics obtained from the three data sets are merely exploratory, but still presented for completeness.

#### 4.2.1 Optimal statistic

We first assess the significance of Hellings-Downs cross-correlations with the optimal statistic (OS; Anholm et al. 2009; Chamberlin et al. 2015) using the `defiant` tool.<sup>2</sup> We specifically use the multiple component OS to measure all three ORFs (Monopole, Dipole, Hellings-Downs) simultaneously, which reduces any bias incurred while measuring one ORF due to the presence of another (Sardesai et al. 2023). To account for CRN and intrinsic pulsar noise parameter uncertainties, we apply noise marginalization to obtain a distribution of the OS (i.e. the NMOS; Vigeland et al. 2018) over 3000 draws from our Bayesian posteriors. For each data set, we then follow Vallisneri et al. (2023) to obtain a Bayesian S/N ( $S/N_{\text{Bayes}}$ ), which can be interpreted as a probability-weighted mean over noise marginalized S/N distribution. For our purposes we assume the GWB S/N null distribution is Gaussian (noting an accurate measure of GW significance mandates use of a GX2 null distribution; see Hazboun

et al. 2023). Our analysis also neglects to account for covariance between pulsar pairs, which naturally arises in pulsar pairs with pulsars in common (Allen & Romano 2023; Johnson et al. 2024). However, this assumption is justified here as we are in the weak S/N regime of the GWB cross-correlations (Sardesai et al. 2023).

We report the  $S/N_{\text{Bayes}}$  values we obtain for Hellings-Downs ORF on the right-most column of Table 3. Using DR2 Lite, We are completely unable to resolve HD correlations from noise, as given by  $S/N_{\text{Bayes}} = 0$ . The combined data sets yield higher, but still relatively low values of  $S/N_{\text{Bayes}} = 0.79$  from Full DR2 and  $S/N_{\text{Bayes}} = 0.92$  from EDR2, with EDR2 yielding slightly higher significance despite including fewer pulsars than Full DR2. These results appear to be consistent with expectations from Agazie et al. (2025), which found that the 30 *noisiest* pulsars can be dropped from the NANOGrav 15-yr data set before the Hellings-Downs S/N experiences any appreciable drop (here EDR2 is effectively Full DR2 with the 31 noisiest pulsars dropped). In this case, a slight increase in S/N with EDR2 may imply either excess noise from a set of the 31 remaining pulsars, or it is simply a statistical fluctuation given the low values of  $S/N_{\text{Bayes}}$ .

Fig. 9 further shows the full distributions of the squared amplitude  $\hat{A}^2$  of each correlation signature, computed using the multiple component NMOS from each data set with uncertainty sampling (Gersbach et al. 2025). The Hellings-Downs GWB amplitudes are consistent with each other and the reported values of  $S/N_{\text{Bayes}}$ , with the combined data sets reducing the uncertainty from DR2 Lite. The

<sup>2</sup><https://github.com/GersbachKa/defiant/tree/main>

presence of monopolar or dipolar correlations indicates additional systematic correlated noise (see Section 3.3). Most notable is that DR2 Lite produces a larger monopole than Full DR2, suggesting the presence of additional unmodelled noise in the uncombined data, which is then mitigated via data combination. The monopole amplitude is further reduced in EDR2 from Full DR2, implying a set of the 31 remaining pulsars may be partially responsible for the noise contributing to this monopole. The DR2 Lite and Full DR2 monopoles both correspond to  $S/N \sim 2$  (Antoniadis et al. 2022). This  $S/N$  level is not particularly significant, as it was shown in Agazie et al. (2023a) that monopolar cross-correlations with higher corresponding  $S/N$  may arise frequently in simulations containing solely HD-correlated GWs and intrinsic pulsar noise. As such, it is possible these monopole measurements are statistical fluctuations, as opposed to the result of a real systematic such as a clock error. We leave a deeper analysis to understand the emergence and nature of monopolar cross-correlations in PTA data sets for future work. The dipole amplitude is highly consistent with zero in each case, which makes sense as the Solar system ephemeris version (DE436) used in this analysis is fixed, independently of how many TOAs are combined. However, the constraints on the amplitude of the dipolar correlations improve as data is combined.

#### 4.2.2 Bayes factors for Hellings-Downs correlations over common noise

IPTA DR2 does not contain enough data to detect the Hellings-Downs curve: Antoniadis et al. (2022) reports a Bayes factor  $\log_{10} \mathcal{B}_{\text{CRN}}^{\text{HD}} \sim 0.3$  ( $\mathcal{B}_{\text{CRN}}^{\text{HD}} \sim 2$ ) for an Hellings-Downs-correlated model versus the autocorrelated CRN model. Nevertheless, this value still indicates a modest preference for an Hellings-Downs cross-correlated model in IPTA DR2. As such, it is valuable to compare the Bayes factors for Hellings-Downs correlations obtained from each data set to find out if DR2 Lite can resolve Hellings-Downs correlations at the same level as the Full DR2.

We report the Bayes factors we measure using each data set in the middle column of Table 3. To check consistency, we report the Bayes factors and sampling uncertainties measured using likelihood reweighting (Hourihane et al. 2023) and using product-space sampling, also known as the HyperModel in enterprise\_extensions (Hee et al. 2016; Johnson et al. 2024). We estimate the uncertainties on the Bayes factors using the effective sample size for reweighting (Hourihane et al. 2023), and using bootstrapping methods for the HyperModel. The two methods agree on the Bayes factor estimates within sampling uncertainties in all cases except for DR2 Lite. However, it is likely that the uncertainty using the HyperModel is underestimated, as the model switch parameter usually has a large autocorrelation length. In comparison to Full DR2 ( $\mathcal{B} \sim 1.4$ ), the cross-correlated model is less favoured using DR2 Lite ( $\mathcal{B} \sim 0.6$ ) and slightly more favoured using EDR2 ( $\mathcal{B} \sim 2.8$ ; Table 3). Although it is unexpected and interesting that EDR2 should return the highest Bayes factor, this appears to be consistent with the OS analysis, especially given we are not considering the presence of a monopole. Overall, even though the Bayes factors are all  $\mathcal{O}(1)$ , these results are a promising sign regarding the potential for combined data sets to improve measurements of cross-correlations over Lite data sets.

### 4.3 Single pulsar noise

We next compare how the characterization of pulsar noise changes whether we use Full DR2 or DR2 Lite. Out of the 53 pulsars in IPTA

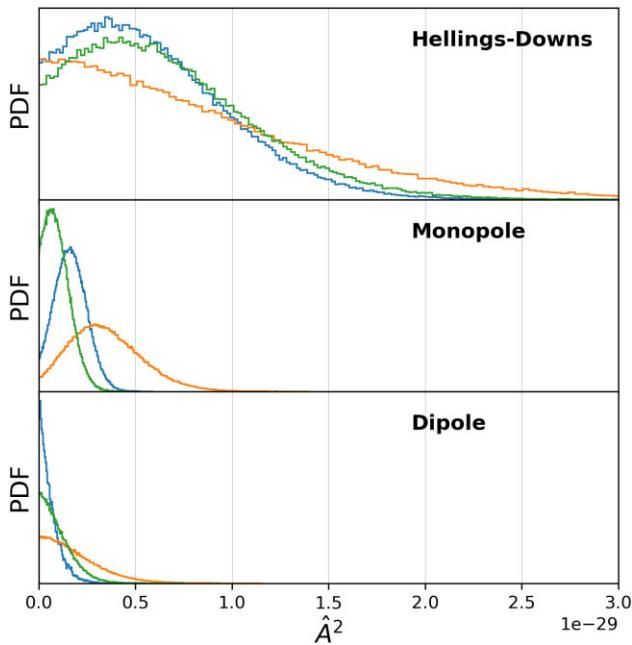
DR2, 24 have timing data from 2 or more PTAs, while the remaining pulsars' data are the same in Full DR2 and DR2 Lite. Therefore, we focus on how the noise properties, primarily red noise and chromatic noise, compare across these 24 pulsars. We do not assess the changes to the white noise or timing model parameters, but we acknowledge these parameters also play a large role in the total characterization of the pulsar.

First, we examine how many new noise processes are detected using Full DR2 versus DR2 Lite. Table 4 shows the Bayes factors measured for each noise process and pulsar using DR2 Lite and Full DR2. These are estimated using the Savage-Dickey density ratio applied to the amplitude parameter of each noise process, e.g.  $\mathcal{B}_0^{\text{PLRN}}$  in PSR J0030 + 0451 is the Bayes factor for a model with DM noise and red noise versus a model with DM noise only. We compare the Bayes factors for red noise, DM noise, higher-order chromatic noise, and time-dependent perturbations to the solar wind density. Bold parameters indicate the noise process is measured with  $\mathcal{B} > 10$ , and where there are no posterior samples in the tail, we place a lower limit of  $\mathcal{B} > 1000$ . Dashes indicate the process was not included in that pulsar's noise model. While these Bayes factors all depend on our choice for how to construct the prior, we are applying the same model and priors to each data set; therefore, the prior-bias becomes less important for the purpose of performing a comparison.

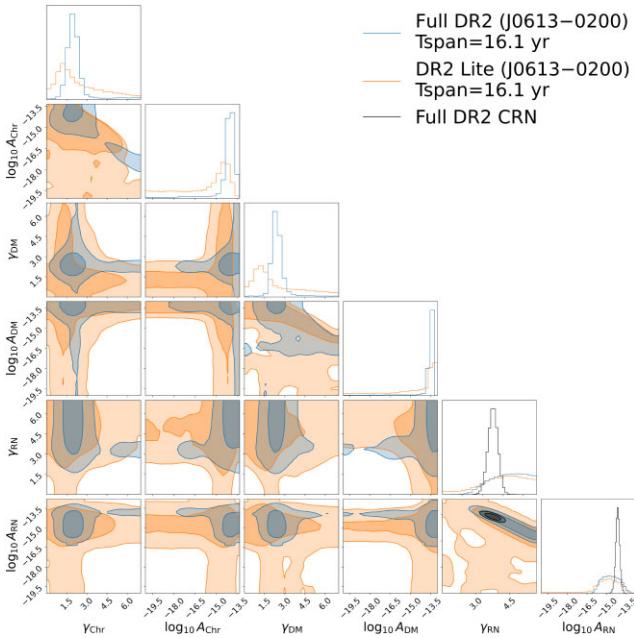
Using the threshold  $\mathcal{B} > 10$  to indicate a probable detection of the noise process, we find using Table 4 that using DR2 Lite, 2/24 pulsars detect achromatic red noise, 5/24 detect DM noise, 1/24 pulsars detect higher-order chromatic noise, and 0/24 detect solar wind density variations. Meanwhile, using Full DR2 we find 12/24 pulsars detect achromatic red noise, 16/24 detect DM noise, and 5/24 pulsars detect higher-order chromatic noise, indicating 10 new detections of red noise, 11 new detections of DM noise, and 4 new detections of chromatic noise in the combined data. 2/17 pulsars newly detect solar wind density variations. Overall, 16/24 pulsars newly detect a noise process that was not detected using DR2 Lite.

Additionally, we find some of these noise processes are not detected in the Lite data set specifically because of source confusion. Bayes factors marked with asterisks in Table 4 signify that there is a strong case of parameter covariance between two or more processes in the model, such that the removal of one process from the model increases the detection significance of the other, and vice versa. In other words, one or more processes are favoured by the data, but the data cannot distinguish which one it is (Lentati et al. 2016; EPTA Collaboration 2023b; Ferranti et al. 2025 each discuss this effect in more depth). We observed this behaviour in 8 pulsars using DR2 Lite. For PSRs J1600–3053 and J1643–1224 the source confusion is between DM and chromatic noise. For PSRs J1022+1001, J1024–0719, J1744–1134, J1918–0642, and J2145–0750, the source confusion is between DM and achromatic red noise. For PSR J0613–0200, the data prefer either to include only DM noise, or to include both achromatic and higher-order chromatic noise. This behaviour is detailed more clearly for PSR J0613–0200 in Fig. 10. In all of these cases, using Full DR2 results in strong measurement of 1 or more of these noise processes, resolving the source confusion. Additionally, for PSR J1738+0333 there is no evidence of achromatic red noise or DM noise in DR2 Lite, but using Full DR2 it has entered into the source confusion regime of the two signals.

We further find from comparing the chromatic, DM, and achromatic red noise parameters from the single-pulsar noise analyses of all 24 pulsars that the changes in the noise parameters going from DR2 Lite to Full DR2 fall under three general categories:



**Figure 9.** Full distribution of  $\hat{A}^2$  of Hellings-Downs, monopolar, and dipolar correlations, computed from the noise marginalized multiple component optimal statistic with uncertainty sampling for Full DR2 (blue), DR2 Lite (orange), and EDR2 (green).



**Figure 10.** Comparison of noise parameter posteriors for chromatic, DM, and red noise from the single pulsar noise analysis of PSR J0613-0200 using DR2 Lite (orange) and using Full DR2 (blue). Panels show the amplitude and spectral index parameters for each noise process. Overplotted in the achromatic red noise panels are the CRN parameters from the analysis of Full DR2, as this noise is a subset of the total achromatic red noise in each pulsar. Contours enclose 68 per cent and 95 per cent 2D Bayesian credible intervals.

**(i) Consistency and improved constraints:** In this case, the noise parameter posteriors measured using Full DR2 are more constrained than the posteriors measured using DR2 Lite, but both sets of posteriors are consistent with one another. This demonstrates the expected effect that adding more data results in more constrained posteriors. The majority of pulsars (15 out of 24) best-fitting into this category: PSRs J0030+0451, J0613-0200, J1012+5307, J1022+1001, J1024-0719, J1455-3330, J1600-3053, J1640+2224, J1738+0333, J1744-1134, J1857+0943, J1918-0642, J2010-1323, J2145-0750, J2317 + 1439.

**(ii) Inconsistency and improved constraints** In this case, using the combined data changes the noise parameter posteriors such that there is noticeable tension between DR2 Lite and Full DR2 in the posteriors for one or more parameters. This could possibly arise if we have model misspecification, or if additional unmitigated chromatic noise enters into the data combination. Otherwise, there may be more complicated interactions between the different model components than expected. Three out of twenty-four pulsars most cleanly fit into this category: PSRs J1643-1224, J1713 + 0747, and J1909-3744.

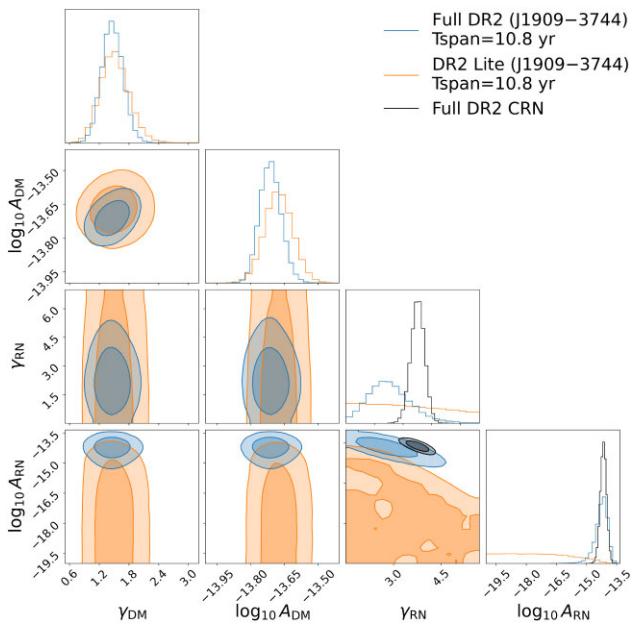
**(iii) Consistency and similar constraints:** In this case, the posterior distributions are very similar and consistent with one another. This indicates the full data combination is not significantly improving noise characterization at the single pulsar level. Six out of twenty-four pulsars best-fitting into this category: PSRs J1730-2304, J1853+1303, J1910+1256, J1939+2134, J1955 + 2908, and J2124-3358.

These categories summarize the effects we observe here of data combination of noise parameter characterization, though some pulsars also toe the line between these categories.

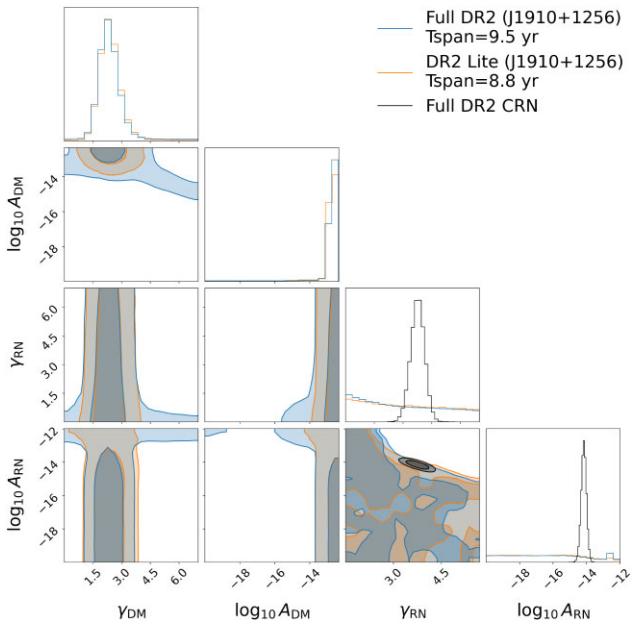
We next follow up with closer examinations of three pulsars, PSRs J0613-0200, J1909-3744, and J1910 + 1256, which each serve as an illustrative case from each category. Figs 10–12 compare their noise parameter distributions obtained using DR2 Lite (orange) and using Full DR2 (blue). We also show the CRN parameters from Full DR2 overplotted over the total achromatic red noise parameters in each figure.

#### 4.3.1 PSR J0613-0200| Consistency and improved constraints

Fig. 10 shows that the Lite data set for PSR J0613-0200 is not sufficient to distinguish the different sources of noise from each other. The 1D posteriors over the DM and red noise amplitudes each have long tails, with posterior support near  $\log_{10} A_{RN} = -20$  and  $\log_{10} A_{DM} = -20$ , indicating the signals are not detected. However, the 2D posterior over both parameters shows a deficiency of samples where both  $\log_{10} A_{RN} = -20$  and  $\log_{10} A_{DM} = -20$ ; i.e. the data does favour inclusion of at least one signal with high significance, but they cannot distinguish which. PSR J0613-0200 also includes higher-order chromatic noise in its model, and this same effect also occurs between  $\log_{10} A_{DM}$  and  $\log_{10} A_{Chr}$ . This means in this case, a side effect of including the higher-order chromatic noise is to slightly *increase* the detection significance for achromatic red noise, rather than to decrease it. This source confusion effect does not occur directly between  $\log_{10} A_{Chr}$  and  $\log_{10} A_{RN}$ , as there are still samples in the 2D region where  $\log_{10} A_{RN} = -20$  and  $\log_{10} A_{DM} = -20$  (instead the L-shape of the 2D 95 per cent credible interval in Fig. 10 is just a product of their 1D posteriors). In contrast, using Full DR2 allows much more precise measurements of each of these processes, and they are better distinguished from one another. These improvements likely result from the improved data cadence



**Figure 11.** Comparison of noise parameter posteriors for DM and red noise from the single pulsar noise analysis of PSR J1909–3744 using DR2 Lite (orange) and using Full DR2 (blue). See Fig. 10 caption for more details.



**Figure 12.** Comparison of noise parameter posteriors for DM and red noise from the single pulsar noise analysis of PSR J1910 + 1256 using DR2 Lite (orange) and using Full DR2 (blue). See Fig. 10 caption for more details.

and radio frequency bandwidth achieved by data combination for this pulsar.

PSR J0613–0200 also shows the strongest preference for the CRN out of the pulsars in EDR2 based on the DFs from Section 4.1.3, as well as the original IPTA DR2 analysis (Antoniadis et al. 2022). This is demonstrated in Fig. 10 by the excellent overlap between the achromatic red noise parameters from the single-pulsar analysis and the CRN parameters from the full-PTA analysis using Full DR2. Meanwhile, the red noise parameters measured using DR2 Lite are

much less constrained, indicating weaker evidence for a detection of red noise as well as a much higher upper limit on red noise. In total, the improved characterization of pulsar noise resulting from the data combination directly translate here to improved measurement of the CRN. We can infer a similar story is at play for several of the other pulsars with improved constraints, particularly those pulsars in Table 4 with covariances recorded between noise parameters, as well as those that show improvements to their DFs in Fig. 7. For other pulsars, such as PSR J2317+1439, the data combination improves their characterization of DM noise but not achromatic red noise.

#### 4.3.2 PSR J1909–3744| Inconsistency & improved constraints

Unlike the case for PSR J0613–0200, Fig. 11 shows that the achromatic red noise distributions obtained using DR2 Lite and Full DR2 are in tension with each other. Using DR2 Lite, no red noise is detected in PSR J1909–3744; however, the red noise detected using Full DR2 lies above the expected region using DR2 Lite (i.e. the 68 per cent 2D credible regions do not overlap). This makes it appear that a red noise process has emerged in the combined data where there was none previously in DR2 Lite. There is no difference in observation timespan between DR2 Lite and Full DR2 for this pulsar, so non-stationary noise is unlikely to cause this. Furthermore, Fig. 11 shows both data sets detect DM variations with similar characteristics in PSR J1909–3744, which makes the possibility of *new* chromatic noise entering into the combined data set seem unlikely. However, the DM noise parameters are measured more precisely using the fully-combined data. Improving the precision on DM variations as a result of improved cadence and radio frequency coverage (shown in Fig. 2) may have the effect of helping to uncover the red noise present in the fully-combined data. This effect only makes sense in tandem with the  $\log_{10}$ -uniform priors used on  $A_{RN}$ , which heavily downweight the presence of noise in the data set. Another possibility is that some level of model misspecification is at play and resulting in inconsistent noise properties across the two data sets.

PSR J1909–3744 is the pulsar in the most tension with the CRN using DR2 Lite, with the lowest DF out of all pulsars (Fig. 7). The CRN parameters from Full DR2 overplotted in Fig. 11 help explain this – they are well above the 95 per cent Bayesian credible interval of the total achromatic red noise measured using the Lite data set under the  $\log_{10}$ -uniform prior. Meanwhile, using EDR2, PSR J1909–3744 is more agnostic to the CRN measurement, indicated in Fig. 11 by the modest overlap between the achromatic red noise using Full DR2 and the CRN. These effects are also evident in the corner plots for PSRs J1713+747 and J1744–1134 (not shown): the achromatic red noise measured using DR2 Lite lies below the level of the red noise measured using Full DR2.

#### 4.3.3 PSR J1910+1256| Consistency & similar constraints

PSR J1910+1256 is the last unique case we cover here. Unlike the previous cases, Fig. 12 shows the posteriors measured for PSR J1910+1256 using DR2 Lite are similar to the posteriors measured using Full DR2, with nearly identical red noise and very similar DM noise. Thus, there apparently exist cases where data combination does not improve noise characterization at the single-pulsar level. The largest factor playing into this may simply be that not enough new data are added into the combination. Indeed, PSR J1910+1256 has one of the lowest FoMs out of the multi-PTA pulsars (Fig. 1), also has a DF  $\sim 1$  (Fig. 7), indicating its combined data set is not yet advanced enough to measure the CRN. While a future combination

may likely yield the emergence of a red noise process in this and similar pulsars, it is also important to consider that pulsar sensitivity may eventually saturate, and therefore reduce the effect of combined data. For example, even with increasing telescope sensitivity, some pulsar's will eventually become jitter-noise limited. However, it does not seem we have reached this regime for the majority of pulsars (Lam et al. 2019). Other pulsars, such as PSR J1939+2134, may be dominated by very strong intrinsic red noise processes, explaining why PSR J1939+2134's noise is well-measured in both DR2 Lite and Full DR2.

#### 4.4 Ensemble noise properties

For our final analysis, we use the framework of hierarchical Bayesian modelling (Loredo & Hendry 2019; Thrane & Talbot 2019; van Haasteren 2024) to estimate the ensemble noise properties of intrinsic red noise and DM variations from each data set (DR2 Lite, EDR2, and Full DR2), following closely the procedures described in Goncharov & Sardana (2025). The ensemble properties are encoded in the distribution  $p(\boldsymbol{\theta}|\mathcal{M}_\Lambda)$ , where we define  $\boldsymbol{\theta}$  as our noise parameters  $\boldsymbol{\theta} = \{\log_{10} A_{\text{RN}}, \gamma_{\text{RN}}, \log_{10} A_{\text{DM}}, \gamma_{\text{DM}}\}$ .  $\mathcal{M}_\Lambda$  designates the hierarchical model with hyperparameters  $\Lambda$ , which are learned from the data  $\delta t$ , while  $\mathcal{M}_\odot$  designates the original model using uninformative priors. For the current analysis, the ensemble noise distribution  $p(\boldsymbol{\theta}|\mathcal{M}_\Lambda)$  is neither a prior nor a posterior on pulsar noise parameters. Rather, it is a posterior predictive distribution for noise parameters, as it is informed by the population of millisecond pulsars in the current data set. For an independent analysis,  $p(\boldsymbol{\theta}|\mathcal{M}_\Lambda)$  could be used as a population-informed prior distribution on the noise parameters.

We estimate the ensemble properties following Goncharov & Sardana (2025) by first inferring the hyperparameters  $\Lambda$  from single pulsar analyses using the marginalized likelihood obtained from importance sampling (Thrane & Talbot 2019):

$$\mathcal{L}(\delta t|\Lambda, \mathcal{M}_\Lambda) = \prod_i^{N_{\text{psr}}} \mathcal{Z}(\delta t_i|\mathcal{M}_\odot) \int \frac{\pi(\boldsymbol{\theta}_i|\Lambda, \mathcal{M}_\Lambda)}{\pi(\boldsymbol{\theta}_i|\mathcal{M}_\odot)} \times \mathcal{P}(\boldsymbol{\theta}_i|\delta t_i, \mathcal{M}_\odot) d\boldsymbol{\theta}_i \quad (14)$$

$$\cong \prod_i^{N_{\text{psr}}} \frac{\mathcal{Z}(\delta t_i|\mathcal{M}_\odot)}{n_i} \sum_k^{n_i} \frac{\pi(\boldsymbol{\theta}_i^k|\Lambda, \mathcal{M}_\Lambda)}{\pi(\boldsymbol{\theta}_i^k|\mathcal{M}_\odot)}, \quad (15)$$

where  $N_{\text{psr}}$  is the number of pulsars,  $n_i$  is the number of posterior samples obtained from the single pulsar  $i$ 's noise analysis, and  $\mathcal{Z}(\delta t_i|\mathcal{M}_\odot)$  is the evidence from the single pulsar  $i$ 's noise analysis (the evidences may be treated as an unknown normalization for the purpose of parameter estimation). In equation (15), each pulsar's integral is estimated by iterating over the  $n_i$  posterior samples (i.e.  $\boldsymbol{\theta}_i^k$  is the  $k$ th posterior sample for the  $i$ th pulsar). As suggested by Goncharov & Sardana (2025), we attempt to separate intrinsic red noise contributions from the GWB by drawing red and DM noise samples from the single pulsar posteriors obtained from our factorized likelihood analyses, which each contained an additional red noise term in each pulsar at fixed  $\gamma = 13/3$  which provides a channel to separate the GWB from intrinsic noise. After defining a suitable hyperprior  $\pi(\Lambda|\mathcal{M}_\Lambda)$ , we can obtain samples over the posterior  $\mathcal{P}(\Lambda|\delta t, \mathcal{M}_\Lambda)$  from stochastic sampling, and numerically marginalize over  $\Lambda$  to obtain the ensemble noise distribution (Goncharov & Sardana 2025)

$$p(\boldsymbol{\theta}|\mathcal{M}_\Lambda) = \int \pi(\boldsymbol{\theta}|\Lambda, \mathcal{M}_\Lambda) \mathcal{P}(\Lambda|\delta t, \mathcal{M}_\Lambda) d\Lambda \quad (16)$$

$$\cong \frac{1}{n_p} \sum_k^{n_p} \pi(\boldsymbol{\theta}|\Lambda_k, \mathcal{M}_\Lambda), \quad (17)$$

where the second line suggests we concatenate samples from the prior  $\pi(\boldsymbol{\theta}|\Lambda_k, \mathcal{M}_\Lambda)$  over  $n_p$   $\Lambda$  samples drawn from the posterior  $\mathcal{P}(\Lambda|\delta t, \mathcal{M}_\Lambda)$ . Equation (16) shows how  $p(\boldsymbol{\theta}|\mathcal{M}_\Lambda)$  implicitly depends on the data  $\delta t$ , and as such it should be interpreted as a posterior predictive distribution for pulsar noise parameters rather than a true prior on  $\boldsymbol{\theta}$ .

We are left with some freedom to define the hyperparameters  $\Lambda$ , the functional form of  $\pi(\boldsymbol{\theta}|\Lambda, \mathcal{M}_\Lambda)$ , and the hyperprior  $\pi(\Lambda|\mathcal{M}_\Lambda)$ . We choose to keep the uniform priors on  $\boldsymbol{\theta}$  and infer the min and max ranges of the prior as our hyperparameters  $\Lambda$ , such that

$$\pi(\theta_j|\Lambda_{\text{max},j}, \Lambda_{\text{min},j}) = \begin{cases} \frac{1}{\Lambda_{\text{max},j} - \Lambda_{\text{min},j}} & \text{if } \Lambda_{\text{min},j} < \theta_j < \Lambda_{\text{max},j}, \\ 0 & \text{elsewhere,} \end{cases} \quad (18)$$

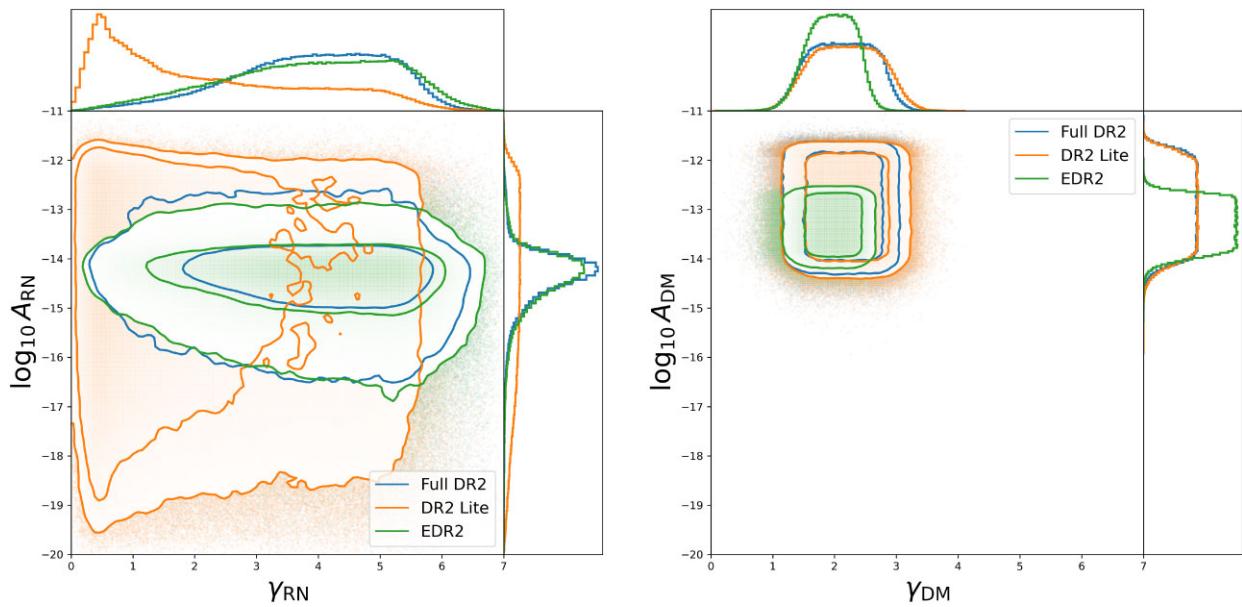
for each of our four noise parameters denoted by  $j$ . This choice is based on Goncharov & Sardana (2025), who found using EPTA DR2 that this uniform distribution model was preferred with a higher Bayes factor than alternative models using a normal distribution or a mixture of normal and uniform distributions. We enforce the constraint  $\Lambda_{\text{min},j} < \Lambda_{\text{max},j}$  by defining our hyperpriors to draw from the distributions

$$p(\Lambda_{\text{min},j}) = \frac{2(u_j - \Lambda_{\text{min},j})}{u_j - l_j} \quad \{l_j < \Lambda_{\text{min},j} < u_j\}, \quad (19)$$

$$p(\Lambda_{\text{max},j}|\Lambda_{\text{min},j}) = \frac{1}{u_j - \Lambda_{\text{min},j}} \quad \{\Lambda_{\text{min},j} < \Lambda_{\text{max},j} < u_j\}, \quad (20)$$

where  $u_j$  and  $l_j$  are the original upper and lower bounds on the red/DM noise parameters (Section 3.2).

We use NUMPYRO (Bingham et al. 2019; Phan, Pradhan & Jankowiak 2019) to define our hierarchical Bayesian model and JAXNS (Albert 2020) to infer the posterior distribution  $\mathcal{P}(\Lambda|\delta t, \mathcal{M}_\Lambda)$  using nested sampling. Fig. 13 shows the resulting ensemble distributions of the parameters of our two noise processes,  $p(\log_{10} A_{\text{RN}}, \gamma_{\text{RN}}|\mathcal{M}_\Lambda)$  and  $p(\log_{10} A_{\text{DM}}, \gamma_{\text{DM}}|\mathcal{M}_\Lambda)$  after numerical marginalization over the hyperparameters  $\Lambda$ . The left hand side shows that DR2 Lite does not produce strong constraints on the intrinsic red noise properties of the ensemble, whereas EDR2 produces equal constraints on the ensemble noise properties as Full DR2. This is consistent in particular with the single pulsar results table 4 which shows only 2 pulsars out of a subset of 24 have strongly detected red noise using DR2 Lite (the list increases to 5 out of 53 once including PSRs J0437–4715, J0621+1002, and J1824–2452A), while Full DR2 is capable of constraining red noise in the majority of pulsars. Meanwhile, the right-hand side of Fig. 13 shows that DR2 Lite can be used to place similar levels of constraints on the ensemble properties of DM variations as Full DR2, even despite detecting DM noise in fewer pulsars than Full DR2 (but still detecting more DM noise processes than red noise processes; Table 4). The distributions are centred near  $\gamma_{\text{DM}} \sim 2$ , but display errors consistent with the value  $\gamma_{\text{DM}} = 8/3$  expected for DM variations from Kolmogorov turbulence (Keith et al. 2013). EDR2 in the meantime displays a tighter distribution of DM noise properties, centred near lower amplitude and spectral index. This is most likely because EDR2 contains only 22 pulsars, and neither PSRs J1721–2457 and J1903+0327, two high-DM pulsars with the highest DM noise amplitudes in IPTA DR2, are included among the 22.



**Figure 13.** Distributions encoding the ensemble noise properties of the IPTA DR2 pulsars as compared whether using DR2 Lite (53 pulsars), EDR2 (22 pulsars), or Full DR2 (53 pulsars). *Left:* Intrinsic red noise properties of the ensemble (decoupled from a separate  $\gamma = 13/3$  process in each pulsar) are substantially broader using DR2 Lite as opposed to EDR2 and Full DR2, which each place equally competitive constraints. *Right:* DM noise properties are tightly constrained using all three data sets, with the EDR2 analysis producing an overly-tight distribution due to the inclusion of fewer pulsars.

Overall, this analysis complements and reinforces our results throughout the rest of this work that DR2 Lite is not as informative as the combined data sets, while the best subset of combined pulsars comprising EDR2 produce comparable results to a full-combination of all pulsars.

## 5 SUMMARY AND DISCUSSION

### 5.1 Summary of the Lite method

The ‘Lite’ method is a novel and computationally efficient approach to create joint-PTA data sets for GW searches. Full data combination, while essential for maximizing GW sensitivity, is a meticulous and resource-intensive process. The Lite method circumvents these challenges by focusing on individual pulsar data sets, enabling quick searches for different GW signals using the latest data. Since Lite data sets can be created on-the-fly using existing PTA data, they are particularly valuable in scenarios where timely results are required, such as when evaluating newly acquired data or gauging the potential of preliminary observations.

By using a FoM to maximize GW sensitivity, Lite data sets represent the optimal combination of single-PTA pulsar data streams that can be achieved prior to the formal combination of all pulsars. The FoM used here comes from the analytic scaling law for the GWB S/N in the intermediate-signal regime, as presented in Siemens et al. (2013). This serves as a PTA-agnostic and statistically robust metric for selecting the most sensitive pulsar data streams from each PTA, while ensuring that the initial analysis retains focus on pulsars that contribute the most to detecting the GW signal of interest, i.e. the GWB. Furthermore, the Lite method allows the flexibility to use different FoMs to construct Lite data sets optimized for different GW signals, such as continuous GWs. We emphasize, however, that the FoMs we use here also make unrealistic assumptions about the data, such as ignoring the effects of uneven data sampling, the presence of chromatic noise, or loud intrinsic red noise. Radio-frequency band

coverage in particular should be an important factor for selecting maximally sensitive single-PTA data sets (Sosa Fiscella et al. 2024; Ferranti et al. 2025). One way to overcome these limitations and create more realistic FoMs in the future might be to numerically compute a FoM using sensitivity curves (Hazboun, Romano & Smith 2019; Baier, Hazboun & Romano 2025). Future metrics might also take the narrow bandwidths into account by adjusting their errorbars using later measurements of DM in the pulsar (Sosa Fiscella et al. 2024). An additional improvement would be to account for sky separation between pulsars pairs when computing the S/N for a GWB, as was done in Speri et al. (2023).

### 5.2 Summary of results using IPTA DR2

To test the Lite method, we create a Lite version of IPTA DR2 (Perera et al. 2019), which we call DR2 Lite. We first carry out a GWB search on DR2 Lite, next comparing the results with a truncated version of Full DR2 using the 22 most sensitive pulsars, which we call EDR2, and finally comparing with Full DR2 (first analysed in Antoniadis et al. 2022). Each version of the data set keeps all single-frequency and legacy data.

We measure a significant CRN process in all versions of IPTA DR2, including DR2 Lite. This provides a proof of concept that a Lite data set may improve upon the individual data sets provided by the regional PTAs while also providing a sneak peek at the signals which may be detected in a fully-combined data set. However, the combined data sets still remain superior. Using the Factorized Likelihood, we show that the detection significance of the CRN improves by orders of magnitude progressing through each stage of the combination process. Data combination also improves the precision of CRN parameter inference. Although the evidence for Hellings-Downs correlations in IPTA DR2 is weak, we tentatively find using both a Bayesian and an OS analysis that the combined data contain more support for Hellings-Downs correlations than DR2 Lite. Finally, across the board EDR2 produces nearly identical cross-

correlation statistics and spectral characterization results as Full DR2, demonstrating that indeed using the FoM or similar statistic to inform the order in which to combine pulsars may yield valuable combined data set at the intermediate stage of data combination, before all pulsars have been combined.

Alongside the above results, we found the amplitude distribution of the CRN is shifted to higher values using DR2 Lite than using EDR2 or Full DR2, suggesting unmitigated intrinsic noise is present in some of the DR2 Lite pulsars and leaking into the common channel. This conclusion is reaffirmed using various ‘dropout’ analyses that show the CRN amplitude using DR2 Lite is strongly dependent on individual pulsars. The multiple component OS analysis also finds a larger amplitude of monopolar correlations in DR2 Lite. This monopole could be partially responsible or connected to the larger CRN amplitude, or it may just be a statistical fluctuation. By using combined data, the amplitude distributions of both the CRN and the monopole shift to lower values, which shows that data combination is capable of mitigating these systematic errors with no required knowledge of their source.

While these results are encouraging, we have not fully stress-tested the Lite method or forecasted its potential for future data sets – this would require analyses of numerous detailed simulations of the data combination process, which is beyond the scope of this work. Another caveat is we do not omit or otherwise account for legacy or single-frequency data in IPTA DR2 from any version of the analysis. As shown by EPTA Collaboration (2023c); Ferranti et al. (2025), including legacy data in the analysis helps to constrain the spectrum but is less useful for measuring the Hellings-Downs curve.

Finally, the effects of data combination are also noticeable on the single pulsar level. After examining 24 pulsars in DR2 Lite which have multi-PTA data in Full DR2, we detect at least one new intrinsic noise process using Full DR2 that we could not detect using DR2 Lite in 16 out of 24 pulsars. These improvements in noise characterization appear to correlate with improvements in the effective radio frequency bandwidth and data cadence that result from data combination. Meanwhile, DR2 Lite resulted in similar constraints on pulsar noise for only 6 out of 24 pulsars. In most of these cases, it appears the data combination did not improve band coverage at low frequencies enough to improve constraints on DM variations, and/or the Lite data set already contained the majority of the TOAs in Full DR2, which likely dominated the statistics. These conditions are unlikely to hold true for many pulsars in the next IPTA data set, as numerous data at low radio frequencies from LOFAR (Stappers et al. 2011), NenuFAR (Zarka et al. 2012), the GMRT (Joshi et al. 2018), and CHIME (CHIME/Pulsar Collaboration 2021) are now becoming available for combination. We should therefore expect to see that the next data combinations will further improve pulsar noise characterization, and by extension, sensitivity to GWs. The reduced number of pulsars with significant detections of noise in DR2 Lite also translate to less informative ensemble distributions of pulsar red noise properties.

### 5.3 Future directions

Looking to the future, the Lite method has a clear role as an intermediate step in the analysis of PTA data sets. This was demonstrated recently in Agazie et al. (2024) where several pseudo-IPTA data sets were composed from the most recent EPTA, NANOGrav, and PPTA data sets using a factorized likelihood approach. While these data sets were not created using a FoM, they are similar in spirit to the Lite method we present here. Agazie et al. (2024) found that adding pulsars to each PTA’s data set using this method consistently results

in a higher GWB S/N and more precise constraints on  $A_{\text{CRN}}$  than what one obtains using each PTA’s individual data release. This further affirms the capability for the Lite method to improve measurements of GW signals. Extrapolating the results of our study comparing IPTA DR2 with its Lite version, we expect the upcoming IPTA data set, IPTA DR3, will further improve GWB spectral characterization and GW detection prospects (Good & International Pulsar Timing Array Team 2023). Furthermore, our analysis of EDR2 reinforces that we can capture a large amount of the information in a fully-combined data set using an intermediate version with fewer pulsars, as expected based on Speri et al. (2023). This suggests the creation and analysis of an IPTA ‘EDR3’ as a way to start reaping the rewards of a fully-combined IPTA DR3 at an earlier time. Though we do not explore this here, the creation of even more optimal joint-PTA data sets should also be possible with a hybrid approach by using combined data for the most sensitive pulsars and using uncombined data for the remaining pulsars.

The Lite analysis may have further unquantified benefits if curating data sets for other nHz GW searches, such as continuous GWs from individual SMBHBs, as it may be used to rapidly improve the sky coverage of the PTA. Of the regional PTAs, three are northern hemisphere (CPTA, EPTA, and NANOGrav), one is quasi-equatorial (InPTA), and two are Southern hemisphere (MPTA and PPTA). The maximum sky coverage achieved by any PTA is approximately 75 per cent due to telescope elevation limits, while for transit telescopes such as CHIME, the sky coverage is considerably less. This sky coverage is perhaps only a secondary concern for the initial studies of the isotropic stochastic GWB, but it is a major concern for efforts to detect and study individual GW sources, as well GWB anisotropy. These gains in sky coverage may be achieved immediately with the creation of Lite data sets.

In conclusion, the Lite method does not replace the need for full data combination, but serves as a powerful exploratory tool for evaluating new data sets. By providing early indicators of GWB sensitivity, the Lite method may help motivate the creation of fully-combined data sets, while the FoM may guide the order in which to combine the pulsar’s data together. As future IPTA data releases incorporate larger and more complex data sets, the Lite method will become increasingly more useful for nHz GW searches as one balances the trade-off between computational efficiency and sensitivity. The Lite method may also complement other future avenues of analysing joint-PTA data sets, such as the Fourier-space combination of posteriors from single-PTA data sets (Valtolina & van Haasteren 2024; Laal et al. 2025), or the fully extended PTA analysis from Agazie et al. (2024) using the OS.

## ACKNOWLEDGEMENTS

This paper is the result of the work of many people and uses data from over three decades of pulsar timing observations. We thank Boris Goncharov for reviewing the manuscript and providing useful suggestions, such as the hierarchical Bayesian analysis, which improved the quality of this work. BL additionally thanks members of the IPTA Gravitational Wave Analysis working group, co-chaired by Nihan Pol, Aurélien Chalumeau, and Paul Baker, for constructive comments and discussions, and to Rutger van Haasteren for providing useful insights on hierarchical modelling. We acknowledge support received from National Science Foundation (NSF) Astronomy and Astrophysics Research Grants (AAG) award number 2009468, and NSF Physics Frontiers Center (PFC) award number 2020265, which supports the NANOGrav project. Part of this work was undertaken as part of the Australian Research Council

Centre of Excellence for Gravitational Wave Discovery (project number CE230100016). The National Radio Astronomy Observatory is a facility of the NSF operated under cooperative agreement by Associated Universities, Inc. The Nançay Radio Observatory is operated by the Paris Observatory, associated to the French Centre National de la Recherche Scientifique (CNRS) and to the Université d'Orléans. We acknowledge financial support from 'Programme National de Cosmologie et Galaxies' (PNCG), and 'Programme National Hautes Energies' (PNHE) funded by CNRS/INSU-IN2P3-INP, the 'Commissariat à l'Énergie Atomique et aux Énergies Alternatives' (CEA) and 'Centre National d'Études Spatiales' (CNES), France. We acknowledge financial support from Agence Nationale de la Recherche (ANR-18-CE31-0015), France. CMFM was supported in part by the NSF under Grant Numbers PHY-2020265 and AST-2106552, and the Flatiron Institute, which is part of the Simons Foundation. JSH acknowledges support from NSF CAREER award number 2339728 and is supported through an Oregon State University start up fund. JA acknowledges support from the European Commission under project ARGOS-CDS (Grant Agreement Number: 101094354). AC acknowledges financial support provided under the European Union's Horizon Europe European Research Council (ERC) Starting Grant 'A Gamma-ray Infrastructure to Advance Gravitational Wave Astrophysics' (GIGA). ZC is supported by the National Natural Science Foundation of China under Grant Number 12405056 and the innovative research group of Hunan Province under Grant Number 2024JJ1006. DD acknowledges the Department of Atomic Energy, Government of India's support through 'Apex Project-Advance Research and Education in Mathematical Sciences' at The Institute of Mathematical Sciences. TD acknowledges support from the NSF PFC award number 1430284. TD was partially supported through the NSF Partnerships for International Research and Education (PIRE) programme award number 0968296. ECF is supported by NASA under award number 80GSFC24M0006. DCG is supported by NSF AAG award number 2406919. SMR is a Canadian Institute for Advanced Research (CIFAR) Fellow and is supported by the NSF PFC award 2020265. GMS acknowledges financial support provided under the European Union's H2020 ERC Consolidator Grant 'Binary Massive Black Hole Astrophysics' (B Massive, Grant Agreement: 818691). JBW is supported by the Major Science and Technology Program of Xinjiang Uygur Autonomous Region (Number 2022A03013-4), the Zhejiang Provincial Natural Science Foundation of China (Number LY23A030001). LZ is supported by the National Natural Science Foundation of China (Grant Number 12103069).

## DATA AVAILABILITY

JUPYTER Notebooks and PYTHON scripts used to produce our figures and our results are freely available at [https://github.com/blarsen10/IPTA\\_DR2\\_analysis](https://github.com/blarsen10/IPTA_DR2_analysis). Correspondence and requests for materials should be addressed to Bjorn Larsen.

## REFERENCES

Afzal A. et al., 2023, *ApJ*, 951, L11  
 Agazie G. et al., 2023a, *ApJ*, 951, L8  
 Agazie G. et al., 2023b, *ApJ*, 951, L9  
 Agazie G. et al., 2023c, *ApJ*, 951, L10  
 Agazie G. et al., 2023d, *ApJ*, 951, L10  
 Agazie G. et al., 2024, *ApJ*, 966, 105  
 Agazie G. et al., 2025, *ApJ*, 978, 168  
 Alam M. F. et al., 2021, *ApJS*, 252, 5  
 Albert J. G., 2020, preprint ([arXiv:2012.15286](https://arxiv.org/abs/2012.15286))  
 Allen B., Romano J. D., 2023, *Phys. Rev. D*, 108, 043026  
 Allen B., Dhurandhar S., Gupta Y., McLaughlin M., Natarajan P., Shannon R. M., Thrane E., Vecchio A., 2023, preprint ([arXiv:2304.04767](https://arxiv.org/abs/2304.04767))  
 Andrews P., Lam M., Dolch T., 2020, Investigating the Impact of Slicing on Fitted Timing Model Parameters, NANOGrav Memo #4  
 Anholm M., Ballmer S., Creighton J. D. E., Price L. R., Siemens X., 2009, *Phys. Rev. D*, 79, 084030  
 Antoniadis J. et al., 2022, *MNRAS*, 510, 4873  
 Arzoumanian Z. et al., 2014, *ApJ*, 794, 141  
 Arzoumanian Z. et al., 2016, *ApJ*, 821, 13  
 Arzoumanian Z. et al., 2020a, *ApJ*, 900, 102  
 Arzoumanian Z. et al., 2020b, *ApJ*, 905, L34  
 Babak S. et al., 2016, *MNRAS*, 455, 1665  
 Baier J. G., Hazboun J. S., Romano J. D., 2025, *Class. Quant. Grav.*, 42, 075008  
 Begelman M. C., Blandford R. D., Rees M. J., 1980, *Nature*, 287, 307  
 Bingham E. et al., 2019, *J. Mach. Learn. Res.*, 20, 28:1  
 CHIME/Pulsar Collaboration et al., 2021, *ApJS*, 255, 5  
 Caballero R. N. et al., 2016, *MNRAS*, 457, 4421  
 Caprini C., Figueroa D. G., 2018, *Class. Quant. Grav.*, 35, 163001  
 Chalumeau A. et al., 2022, *MNRAS*, 509, 5538  
 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  
 Champion D. J. et al., 2010, *ApJ*, 720, L201  
 Chen S. et al., 2021a, *MNRAS*, 508, 4970  
 Chen S. et al., 2021b, *MNRAS*, 508, 4970  
 Chen S. et al., 2021c, *MNRAS*, 508, 4970  
 Cordes J. M., Rickett B. J., 1998, *ApJ*, 507, 846  
 Desvignes G. et al., 2016, *MNRAS*, 458, 3341  
 Detweiler S., 1979, *ApJ*, 234, 1100  
 Dickey J. M., 1971, *Ann. Math. Stat.*, 42 204  
 EPTA Collaboration, 2023a, *A&A*, 678, A48  
 EPTA Collaboration, 2023b, *A&A*, 678, A49  
 EPTA Collaboration, 2023c, *A&A*, 678, A50  
 Ellis J., van Haasteren R., 2017, *jellis18/PTMCMCSampler: Official Release. Zenodo*  
 Ellis J. A., Vallisneri M., Taylor S. R., Baker P. T., 2020, *ENTERPRISE: Enhanced Numerical Toolbox Enabling a Robust Pulsar Inference Suite. Zenodo*  
 FERMI-LAT Collaboration, 2022, *Science*, 376, 521  
 Falxa M. et al., 2023, *MNRAS*, 521, 5077  
 Ferranti I. et al., 2025, *A&A*, 694, A38  
 Gaia Collaboration, 2021, *A&A*, 649, A1  
 Geiger A. et al., 2025, *ApJ*, 986, 191  
 Gersbach K. A., Taylor S. R., Meyers P. M., Romano J. D., 2025, *Phys. Rev. D*, 111, 023027  
 Goncharov B., Sardana S., 2025, *MNRAS*, 537, 3470  
 Goncharov B. et al., 2021a, *MNRAS*, 502, 478  
 Goncharov B. et al., 2021b, *ApJ*, 917, L19  
 Goncharov B. et al., 2022, *ApJ*, 932, L22  
 Goncharov B. et al., 2024, preprint ([arXiv:2409.03627](https://arxiv.org/abs/2409.03627))  
 Good D., International Pulsar Timing Array Team, 2023, in American Astronomical Society Meeting Abstracts. p. 438.02  
 Hazboun J. S., Romano J. D., Smith T. L., 2019, *Phys. Rev. D*, 100, 104028  
 Hazboun J. S. et al., 2020, *ApJ*, 890, 108  
 Hazboun J. S. et al., 2022, *ApJ*, 929, 39  
 Hazboun J. S., Meyers P. M., Romano J. D., Siemens X., Archibald A. M., 2023, *Phys. Rev. D*, 108, 104050  
 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  
 Hemberger D. A., Stinebring D. R., 2008, *ApJ*, 674, L37  
 Hobbs G. et al., 2012, *MNRAS*, 427, 2780  
 Hourihane S., Meyers J., Johnson A., Chatzioannou K., Vallisneri M., 2023, *Phys. Rev. D*, 107, 084045  
 Johnson A. D., Vigeland S. J., Siemens X., Taylor S. R., 2022, *ApJ*, 932, 105

Johnson A. D. et al., 2024, *Phys. Rev. D*, 109, 103012  
 Joshi B. C. et al., 2018, *JA&A*, 39, 51  
 Kaspi V. M., Taylor J. H., Ryba M. F., 1994, *ApJ*, 428, 713  
 Keith M. J. et al., 2013, *MNRAS*, 429, 2161  
 Laal N., Taylor S. R., van Haasteren R., Lamb W. G., Siemens X., 2025, *Phys. Rev. D*, 111, 063067  
 Lam M. T., McLaughlin M. A., Cordes J. M., Chatterjee S., Lazio T. J. W., 2018a, *ApJ*, 861, 12  
 Lam M. T. et al., 2018b, *ApJ*, 861, 132  
 Lam M. T. et al., 2019, *ApJ*, 872, 193  
 Lange J. U., 2023, *MNRAS*, 525, 3181  
 Larsen B. et al., 2024, *ApJ*, 972, 49  
 Lasky P. D. et al., 2016, *Phys. Rev. X*, 6, 011035  
 Lee K. J., 2016, in Qain L., Li D., eds, *ASP Conf. Ser. Vol. 502, Frontiers in Radio Astronomy and FAST Early Sciences Symposium 2015*. Astron. Soc. Pac., San Francisco, p. 19  
 Lentati L., Alexander P., Hobson M. P., Taylor S., Gair J., Balan S. T., van Haasteren R., 2013, *Phys. Rev. D*, 87, 104021  
 Lentati L. et al., 2016, *MNRAS*, 458, 2161  
 Loredo T. J., Hendry M. A., 2019, preprint ([arXiv:1911.12337](https://arxiv.org/abs/1911.12337))  
 Madison D. R., Cordes J. M., Chatterjee S., 2014, *ApJ*, 788, 141  
 Manchester R. N. et al., 2013, *PASA*, 30, e017  
 Meyers P. M., Chatzioannou K., Vallisneri M., Chua A. J. K., 2023, *Phys. Rev. D*, 108, 123008  
 Miles M. T. et al., 2023, *MNRAS*, 519, 3976  
 Mingarelli C. M. F. et al., 2017, *Nat. Astron.*, 1, 886  
 Nijt I. C. et al., 2024, *MNRAS*, 528, 3304  
 Perera B. B. P. et al., 2019, *MNRAS*, 490, 4666  
 Phan D., Pradhan N., Jankowiak M., 2019, preprint ([arXiv:1912.11554](https://arxiv.org/abs/1912.11554))  
 Rajagopal M., Romani R. W., 1995, *ApJ*, 446, 543  
 Ransom S. et al., 2019, in *Bulletin of the American Astronomical Society*. p. 195, preprint ([arXiv:1908.05356](https://arxiv.org/abs/1908.05356))  
 Reardon D. J. et al., 2016, *MNRAS*, 455, 1751  
 Reardon D. J. et al., 2023a, *ApJ*, 951, L6  
 Reardon D. J. et al., 2023b, *ApJ*, 951, L7  
 Reardon D. J. et al., 2024, *ApJ*, 971, L18  
 Saradesai S. C., Vigeland S. J., Gersbach K. A., Taylor S. R., 2023, *Phys. Rev. D*, 108, 124081  
 Sazhin M. V., 1978, *Soviet Ast.*, 22, 36  
 Shannon R. M., Cordes J. M., 2017, *MNRAS*, 464, 2075  
 Shannon R. M. et al., 2015, *Science*, 349, 1522  
 Siemens X., Ellis J., Jenet F., Romano J. D., 2013, *Class. Quant. Grav.*, 30, 224015  
 Sosa Fiscella S. V. et al., 2024, *ApJ*, 966, 95  
 Speri L., Porayko N. K., Falxa M., Chen S., Gair J. R., Sesana A., Taylor S. R., 2023, *MNRAS*, 518, 1802  
 Srivastava A. et al., 2023, *Phys. Rev. D*, 108, 023008  
 Stappers B. W. et al., 2011, *A&A*, 530, A80  
 Susarla S. C. et al., 2024, *A&A*, 692, A18  
 Taylor S. R., 2021, preprint ([arXiv:2105.13270](https://arxiv.org/abs/2105.13270))  
 Taylor S. R., Baker P. T., Hazboun J. S., Simon J., Vigeland S. J., 2021, *enterprise\_extensions. v3.0.3* [https://github.com/nanograv/enterprise\\_extensions](https://github.com/nanograv/enterprise_extensions)  
 Taylor S. R., Simon J., Schult L., Pol N., Lamb W. G., 2022, *Phys. Rev. D*, 105, 084049  
 Thrane E., Talbot C., 2019, *PASA*, 36, e010  
 Turner J. E. et al., 2021, *ApJ*, 917, 10  
 Vallisneri M., Meyers P. M., Chatzioannou K., Chua A. J. K., 2023, *Phys. Rev. D*, 108, 123007  
 Valtolina S., van Haasteren R., 2024, preprint ([arXiv:2412.11894](https://arxiv.org/abs/2412.11894))  
 van Haasteren R., 2024, *ApJS*, 273, 23  
 van Haasteren R., 2025, *MNRAS*, 537, L1  
 van Haasteren R., Levin Y., 2010, *MNRAS*, 401, 2372  
 van Haasteren R., Levin Y., 2013, *MNRAS*, 428, 1147  
 van Haasteren R., Vallisneri M., 2014, *Phys. Rev. D*, 90, 104012  
 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

Volonteri M., Haardt F., Madau P., 2003, *ApJ*, 582, 559  
 You X. P., Hobbs G. B., Coles W. A., Manchester R. N., Han J. L., 2007, *ApJ*, 671, 907  
 Zarka P., Girard J. N., Tagger M., Denis L., 2012, in Boissier S., de Laverny P., Nardetto N., Samadi R., Valls-Gabaud D., Wozniak H., eds, *SF2A-2012: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics*. p. 687  
 Zhu W. W. et al., 2015, *ApJ*, 809, 41  
 Zic A. et al., 2022, *MNRAS*, 516, 410  
 Zic A. et al., 2023a, *PASA*, 40, e049  
 Zic A. et al., 2023b, *PASA*, 40, e049

## APPENDIX A: FIGURES OF MERIT FOR OTHER GRAVITATIONAL WAVE SEARCHES

We highlight here some Figures of Merit that could be used if creating a Lite data set optimized for other types of GW searches than a GWB. For example, a continuous GW from an isolated SMBHB in the white-noise-dominated regime will have the FoM (Arzoumanian et al. 2014; Mingarelli et al. 2017):

$$\text{FoM}_{\text{CGW}} = \left( \frac{T_{\text{obs}} c}{\langle \sigma_{\text{TOA}} \rangle^2} \right)^{1/2}, \quad (\text{A1})$$

where we have used  $c = 1/\langle \Delta t \rangle$ . Here, equation (A1) ignores the contribution of the GW antenna beam pattern, as it may be not known a priori. However, for a targeted SMBHB search (e.g. Arzoumanian et al. 2020a), the contribution of the antenna beam may be included. For a GW burst with memory, we have (van Haasteren & Levin 2010; Madison, Cordes & Chatterjee 2014)

$$\text{FoM}_{\text{BwM}} = \left( \frac{T_{\text{obs}}^3 c}{\langle \sigma_{\text{TOA}} \rangle^2} \right)^{1/2}. \quad (\text{A2})$$

$\text{FoM}_{\text{BwM}}$  and  $\text{FoM}_{\text{CGW}}$  depend more strongly than  $\text{FoM}_{\text{GWB}}$  on  $c/\langle \sigma_{\text{TOA}} \rangle^2$ , and  $\text{FoM}_{\text{BwM}}$  depends more strongly than  $\text{FoM}_{\text{CGW}}$  on  $T_{\text{obs}}$ . This suggests different Lite data sets to be curated for different GW signals. Additional FoMs can be constructed for any type of signal affecting PTAs so long as a theoretical S/N can be defined.

<sup>1</sup>Department of Physics, Yale University, New Haven, CT 06520, USA

<sup>2</sup>Department of Physics and Astronomy, Widener University, One University Place, Chester, PA 19013, USA

<sup>3</sup>Department of Physics, Oregon State University, Corvallis, OR 97331, USA

<sup>4</sup>Shanghai Astronomical Observatory, Chinese Academy of Sciences, 80 Nandan Road, Shanghai 200030, China

<sup>5</sup>Key Laboratory of Radio Astronomy and Technology (Chinese Academy of Sciences), A20 Datun Road, Chaoyang District, Beijing 100101, China

<sup>6</sup>Department of Physics and Astronomy, Vanderbilt University, 2301 Vanderbilt Place, Nashville, TN 37235, USA

<sup>7</sup>Department of Astrophysical and Planetary Sciences, University of Colorado, Boulder, CO 80309, USA

<sup>8</sup>FORTH Institute of Astrophysics, Max-Planck Institut für Radioastronomie, N. Plastira 100, 70013 Heraklion, Greece

<sup>9</sup>Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, China

<sup>10</sup>ASTRON, Netherlands Institute for Radio Astronomy, Oude Hoogeveensedijk 4, NL-7991 PD Dwingeloo, the Netherlands

<sup>11</sup>Department of Physics and Synergetic Innovation Center for Quantum Effects and Applications, Hunan Normal University, Changsha, Hunan 410081, China

<sup>12</sup>Institute of Interdisciplinary Studies, Hunan Normal University, Changsha, Hunan 410081, China

<sup>13</sup>LPC2E, OSUC, Univ. Orléans, CNRS, CNES, Observatoire de Paris, Université PSL, F-45071 Orléans, France

<sup>14</sup>Observatoire Radioastronomique de Nançay, Observatoire de Paris, Université PSL, Univ Orléans, CNRS, F-18330 Nançay, France

<sup>15</sup>*The Institute of Mathematical Sciences, C.I.T. Campus, Taramani, Chennai 600113, India*

<sup>16</sup>*School of Physics and Astronomy, Monash University, Clayton VIC 3800, Australia*

<sup>17</sup>*Australian Research Council Centre of Excellence for Gravitational Wave Discovery (OzGrav), Clayton VIC 3800, Australia*

<sup>18</sup>*Australia Telescope National Facility, CSIRO, Space and Astronomy, PO Box 76, Epping, NSW 1710, Australia*

<sup>19</sup>*Department of Physics, Hillsdale College, College Street, Hillsdale, MI 49242, USA*

<sup>20</sup>*Department of Physics and Astronomy, University of Nigeria, Nsukka 410105, Enugu, Nigeria*

<sup>21</sup>*Department of Astronomy, University of Maryland, College Park, MD 20742, USA*

<sup>22</sup>*Center for Research and Exploration and Space Studies (CRESST), NASA/GSFC, Greenbelt, MD 20771, USA*

<sup>23</sup>*NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA*

<sup>24</sup>*Department of Physics and Astronomy, University of Montana, 32 Campus Drive, Missoula, MT 59812, USA*

<sup>25</sup>*Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany*

<sup>26</sup>*Department of Physics and Astronomy, Macquarie University, Sydney, NSW 2109, Australia*

<sup>27</sup>*SETI Institute, 339 N Bernardo Ave Suite 200, Mountain View, CA 94043, USA*

<sup>28</sup>*School of Physics and Astronomy, Rochester Institute of Technology, Rochester, NY 14623, USA*

<sup>29</sup>*Laboratory for Multiwavelength Astrophysics, Rochester Institute of Technology, Rochester, NY 14623, USA*

<sup>30</sup>*Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA*

<sup>31</sup>*Department of Physics and Astronomy, West Virginia University, PO Box 6315, Morgantown, WV 26506, USA*

<sup>32</sup>*Department of Physics, Lafayette College, Easton, PA 18042, USA*

<sup>33</sup>*Florida Space Institute, University of Central Florida, 12354 Research Parkway, Orlando, FL 32826, USA*

<sup>34</sup>*IRFU, CEA, Université Paris-Saclay, F-91191 Gif-sur-Yvette, France*

<sup>35</sup>*NRAO, 520 Edgemont Road, Charlottesville, VA 22903, USA*

<sup>36</sup>*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, PO Box 218, Hawthorn, Victoria 3122, Australia*

<sup>37</sup>*CSIRO Scientific Computing, Australian Technology Park, Locked Bag 9013, Alexandria, NSW 1435, Australia*

<sup>38</sup>*Dipartimento di Fisica ‘G. Occhialini’, Università degli Studi di Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy*

<sup>39</sup>*INAF – Osservatorio Astronomico di Cagliari, via della Scienza 5, I-09047 Selargius (CA), Italy*

<sup>40</sup>*INFN, Sezione di Milano-Bicocca, Piazza della Scienza 3, I-20126 Milano, Italy*

<sup>41</sup>*European Space Agency (ESA), European Space Research and Technology Centre (ESTEC), Keplerlaan 1, NL-2201 AZ Noordwijk, the Netherlands*

<sup>42</sup>*Department of Physics, GLA University, Mathura 281406, India*

<sup>43</sup>*Institute of Optoelectronic Technology, Lishui University, Lishui, Zhejiang 323000, China*

<sup>44</sup>*Ruhr University Bochum, Faculty of Physics and Astronomy, Astronomical Institute (AIRUB), D-44780 Bochum, Germany*

<sup>45</sup>*National Astronomical Observatories, Chinese Academy of Sciences, A20 Datun Road, Chaoyang District, Beijing 100101, China*

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