



The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background

Gabriella Agazie¹, Akash Anumalapudi¹, Anne M. Archibald², Zaven Arzoumanian³, Paul T. Baker⁴, Bence Bécsey⁵,
 Laura Blecha⁶, Adam Brazier^{7,8}, Paul R. Brook⁹, Sarah Burke-Spolaor^{10,11}, Rand Burnette⁵, Robin Case⁵,
 Maria Charisi¹², Shami Chatterjee⁷, Katerina Chatziioannou¹³, Belinda D. Cheeseboro^{10,11}, Siyuan Chen¹⁴,
 Tyler Cohen¹⁵, James M. Cordes⁷, Neil J. Cornish¹⁶, Fronefield Crawford¹⁷, H. Thankful Cromartie^{7,10},
 Kathryn Crowter¹⁸, Curt J. Cutler^{13,19}, Megan E. DeCesar²⁰, Dallas DeGan⁵, Paul B. Demorest²¹, Heling Deng⁵,
 Timothy Dolch^{22,23}, Brendan Drachler^{24,25}, Justin A. Ellis²⁶, Elizabeth C. Ferrara^{27,28,29}, William Fiore^{10,11},
 Emmanuel Fonseca^{10,11}, Gabriel E. Freedman¹, Nate Garver-Daniels^{10,11}, Peter A. Gentile^{10,11}, Kyle A. Gersbach¹²,
 Joseph Glaser^{10,11}, Deborah C. Good^{30,31}, Kayhan Gültekin³², Jeffrey S. Hazboun⁵, Sophie Hourihane¹³, Kristina Islo¹,
 Ross J. Jennings^{10,11,73}, Aaron D. Johnson^{1,13}, Megan L. Jones¹, Andrew R. Kaiser^{10,11}, David L. Kaplan¹,
 Luke Zoltan Kelley³³, Matthew Kerr³⁴, Joey S. Key³⁵, Tonia C. Klein¹, Nima Laal⁵, Michael T. Lam^{24,25},
 William G. Lamb¹², T. Joseph W. Lazio¹⁹, Natalia Lewandowska³⁶, Tyson B. Littenberg³⁷, Tingting Liu^{10,11},
 Andrea Lommen³⁸, Duncan R. Lorimer^{10,11}, Jing Luo^{39,71}, Ryan S. Lynch⁴⁰, Chung-Pei Ma^{33,41},
 Dustin R. Madison⁴², Margaret A. Mattson^{10,11}, Alexander McEwen¹, James W. McKee^{43,44}, Maura A. McLaughlin^{10,11},
 Natasha McMann¹², Bradley W. Meyers^{18,45}, Patrick M. Meyers¹³, Chiara M. F. Mingarelli^{30,31,46}, Andrea Mitridate⁴⁷,
 Priyamvada Natarajan^{48,49}, Cherry Ng⁵⁰, David J. Nice⁵¹, Stella Koch Ocker⁷, Ken D. Olum⁵²,
 Timothy T. Pennucci⁵³, Benetge B. P. Perera⁵⁴, Polina Petrov¹², Nihan S. Pol¹², Henri A. Radovan⁵⁵,
 Scott M. Ransom⁵⁶, Paul S. Ray³⁴, Joseph D. Romano⁵⁷, Shashwat C. Sardesai¹, Ann Schmiedekamp⁵⁸,
 Carl Schmiedekamp⁵⁸, Kai Schmitz⁵⁹, Levi Schult¹², Brent J. Shapiro-Albert^{10,11,60}, Xavier Siemens^{1,5},
 Joseph Simon^{61,72}, Magdalena S. Siwek⁶², Ingrid H. Stairs¹⁸, Daniel R. Stinebring⁶³, Kevin Stovall²¹, Jerry P. Sun⁵,
 Abhimanyu Susobhanan¹, Joseph K. Swiggum^{51,72}, Jacob Taylor⁵, Stephen R. Taylor¹², Jacob E. Turner^{10,11},
 Caner Unal^{64,65}, Michele Vallisneri^{13,19}, Rutger van Haasteren⁶⁶, Sarah J. Vigeland¹, Haley M. Wahl^{10,11},
 Qiaohong Wang¹², Caitlin A. Witt^{67,68}, and Olivia Young^{24,25}
 The NANOGrav Collaboration⁶⁹

¹ Center for Gravitation, Cosmology and Astrophysics, Department of Physics, University of Wisconsin–Milwaukee, P.O. Box 413, Milwaukee, WI 53201, USA² Newcastle University, NE1 7RU, UK³ X-Ray Astrophysics Laboratory, NASA Goddard Space Flight Center, Code 662, Greenbelt, MD 20771, USA⁴ Department of Physics and Astronomy, Widener University, One University Place, Chester, PA 19013, USA⁵ Department of Physics, Oregon State University, Corvallis, OR 97331, USA⁶ Physics Department, University of Florida, Gainesville, FL 32611, USA⁷ Cornell Center for Astrophysics and Planetary Science and Department of Astronomy, Cornell University, Ithaca, NY 14853, USA⁸ Cornell Center for Advanced Computing, Cornell University, Ithaca, NY 14853, USA⁹ Institute for Gravitational Wave Astronomy and School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK¹⁰ Department of Physics and Astronomy, West Virginia University, P.O. Box 6315, Morgantown, WV 26506, USA¹¹ Center for Gravitational Waves and Cosmology, West Virginia University, Chestnut Ridge Research Building, Morgantown, WV 26505, USA¹² Department of Physics and Astronomy, Vanderbilt University, 2301 Vanderbilt Place, Nashville, TN 37235, USA¹³ Division of Physics, Mathematics, and Astronomy, California Institute of Technology, Pasadena, CA 91125, USA¹⁴ Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing, 100871, People's Republic of China¹⁵ Department of Physics, New Mexico Institute of Mining and Technology, 801 Leroy Place, Socorro, NM 87801, USA¹⁶ Department of Physics, Montana State University, Bozeman, MT 59717, USA¹⁷ Department of Physics and Astronomy, Franklin & Marshall College, P.O. Box 3003, Lancaster, PA 17604, USA¹⁸ Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada¹⁹ Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA²⁰ George Mason University, resident at the Naval Research Laboratory, Washington, DC 20375, USA²¹ National Radio Astronomy Observatory, 1003 Lopezville Rd., Socorro, NM 87801, USA²² Department of Physics, Hillsdale College, 33 E. College Street, Hillsdale, MI 49242, USA²³ Eureka Scientific, 2452 Delmer Street, Suite 100, Oakland, CA 94602-3017, USA²⁴ School of Physics and Astronomy, Rochester Institute of Technology, Rochester, NY 14623, USA²⁵ Laboratory for Multiwavelength Astrophysics, Rochester Institute of Technology, Rochester, NY 14623, USA²⁶ Bionic Health, 800 Park Offices Drive, Research Triangle Park, NC 27709, USA²⁷ Department of Astronomy, University of Maryland, College Park, MD 20742, USA²⁸ Center for Research and Exploration in Space Science and Technology, NASA/GSFC, Greenbelt, MD 20771, USA²⁹ NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA³⁰ Department of Physics, University of Connecticut, 196 Auditorium Road, U-3046, Storrs, CT 06269-3046, USA³¹ Center for Computational Astrophysics, Flatiron Institute, 162 5th Avenue, New York, NY 10010, USA³² Department of Astronomy and Astrophysics, University of Michigan, Ann Arbor, MI 48109, USA³³ Department of Astronomy, University of California, Berkeley, 501 Campbell Hall #3411, Berkeley, CA 94720, USA³⁴ Space Science Division, Naval Research Laboratory, Washington, DC 20375-5352, USA³⁵ University of Washington Bothell, 18115 Campus Way NE, Bothell, WA 98011, USA³⁶ Department of Physics, State University of New York at Oswego, Oswego, NY, 13126, USA³⁷ NASA Marshall Space Flight Center, Huntsville, AL 35812, USA³⁸ Department of Physics and Astronomy, Haverford College, Haverford, PA 19041, USA³⁹ Department of Astronomy & Astrophysics, University of Toronto, 50 Saint George Street, Toronto, ON M5S 3H4, Canada

- ⁴⁰ Green Bank Observatory, P.O. Box 2, Green Bank, WV 24944, USA
⁴¹ Department of Physics, University of California, Berkeley, CA 94720, USA
⁴² Department of Physics, University of the Pacific, 3601 Pacific Avenue, Stockton, CA 95211, USA
⁴³ E.A. Milne Centre for Astrophysics, University of Hull, Cottingham Road, Kingston-upon-Hull, HU6 7RX, UK
⁴⁴ Centre of Excellence for Data Science, Artificial Intelligence and Modelling (DAIM), University of Hull, Cottingham Road, Kingston-upon-Hull, HU6 7RX, UK
⁴⁵ International Centre for Radio Astronomy Research, Curtin University, Bentley, WA 6102, Australia
⁴⁶ Department of Physics, Yale University, New Haven, CT 06520, USA
⁴⁷ Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, D-22607 Hamburg, Germany
⁴⁸ Department of Astronomy, Yale University, 52 Hillhouse Ave., New Haven, CT 06511, USA
⁴⁹ Black Hole Initiative, Harvard University, 20 Garden Street, Cambridge, MA 02138, USA
⁵⁰ Dunlap Institute for Astronomy and Astrophysics, University of Toronto, 50 St. George St., Toronto, ON M5S 3H4, Canada
⁵¹ Department of Physics, Lafayette College, Easton, PA 18042, USA
⁵² Institute of Cosmology, Department of Physics and Astronomy, Tufts University, Medford, MA 02155, USA
⁵³ Institute of Physics and Astronomy, Eötvös Loránd University, Pázmány P.s. 1/A, 1117 Budapest, Hungary
⁵⁴ Arecibo Observatory, HC3 Box 53995, Arecibo, PR 00612, USA
⁵⁵ Department of Physics, University of Puerto Rico, Mayagüez, PR 00681, USA
⁵⁶ National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, VA 22903, USA
⁵⁷ Department of Physics, Texas Tech University, Box 41051, Lubbock, TX 79409, USA
⁵⁸ Department of Physics, Penn State Abington, Abington, PA 19001, USA
⁵⁹ Institute for Theoretical Physics, University of Münster, D-48149 Münster, Germany
⁶⁰ Giant Army, 915A 17th Ave., Seattle WA 98122, USA
⁶¹ Department of Astrophysical and Planetary Sciences, University of Colorado, Boulder, CO 80309, USA
⁶² Center for Astrophysics, Harvard University, 60 Garden St., Cambridge, MA 02138, USA
⁶³ Department of Physics and Astronomy, Oberlin College, Oberlin, OH 44074, USA
⁶⁴ Department of Physics, Ben-Gurion University of the Negev, Be'er Sheva 84105, Israel
⁶⁵ Feza Gursey Institute, Bogazici University, Kandilli, 34684, Istanbul, Turkey
⁶⁶ Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut), Callinstrasse 38, D-30167, Hannover, Germany
⁶⁷ Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA), Northwestern University, Evanston, IL 60208, USA
⁶⁸ Adler Planetarium, 1300 S. DuSable Lake Shore Dr., Chicago, IL 60605, USA
- Received 2023 May 8; revised 2023 May 31; accepted 2023 June 1; published 2023 June 29

Abstract

We report multiple lines of evidence for a stochastic signal that is correlated among 67 pulsars from the 15 yr pulsar timing data set collected by the North American Nanohertz Observatory for Gravitational Waves. The correlations follow the Hellings–Downs pattern expected for a stochastic gravitational-wave background. The presence of such a gravitational-wave background with a power-law spectrum is favored over a model with only independent pulsar noises with a Bayes factor in excess of 10^{14} , and this same model is favored over an uncorrelated common power-law spectrum model with Bayes factors of 200–1000, depending on spectral modeling choices. We have built a statistical background distribution for the latter Bayes factors using a method that removes interpulsar correlations from our data set, finding $p = 10^{-3}$ ($\approx 3\sigma$) for the observed Bayes factors in the null no-correlation scenario. A frequentist test statistic built directly as a weighted sum of interpulsar correlations yields $p = 5 \times 10^{-5}$ to 1.9×10^{-4} ($\approx 3.5\sigma$ – 4σ). Assuming a fiducial $f^{-2/3}$ characteristic strain spectrum, as appropriate for an ensemble of binary supermassive black hole inspirals, the strain amplitude is $2.4^{+0.7}_{-0.6} \times 10^{-15}$ (median + 90% credible interval) at a reference frequency of 1 yr^{-1} . The inferred gravitational-wave background amplitude and spectrum are consistent with astrophysical expectations for a signal from a population of supermassive black hole binaries, although more exotic cosmological and astrophysical sources cannot be excluded. The observation of Hellings–Downs correlations points to the gravitational-wave origin of this signal.

Unified Astronomy Thesaurus concepts: Gravitational waves (678); Gravitational wave astronomy (675); Millisecond pulsars (1062); Radio pulsars (1353); Supermassive black holes (1663)

1. Introduction

Almost a century had to elapse between Einstein’s prediction of gravitational waves (GWs; Einstein 1916) and their measurement from a coalescing binary of stellar-mass black

holes (Abbott et al. 2016). However, their existence had been confirmed in the late 1970s through measurements of the orbital decay of the Hulse–Taylor binary pulsar (Hulse & Taylor 1975; Taylor et al. 1979). Today, pulsars are again at the forefront of the quest to detect GWs, this time from binary systems of central galactic black holes.

Black holes with masses of 10^5 – $10^{10} M_\odot$ exist at the center of most galaxies and are closely correlated with the global properties of the host, suggesting a symbiotic evolution (Magorrian et al. 1998; McConnell & Ma 2013). Galaxy mergers are the main drivers of hierarchical structure formation over cosmic time (Blumenthal et al. 1984) and lead to the formation of close massive black hole binaries long after the mergers (Begelman et al. 1980; Milosavljević & Merritt 2003). The most massive of these (supermassive black hole binaries

⁶⁹ comments@nanograv.org.

⁷⁰ NASA Hubble Fellowship: Einstein Postdoctoral Fellow.

⁷¹ Deceased.

⁷² NSF Astronomy and Astrophysics Postdoctoral Fellow.

⁷³ NANOGrav Physics Frontiers Center Postdoctoral Fellow.



(SMBHBs), with masses 10^8 – $10^{10} M_\odot$ emit GWs with slowly evolving frequencies, contributing to a noise-like broadband signal in the nHz range (the GW background (GWB); Rajagopal & Romani 1995; Jaffe & Backer 2003; Wyithe & Loeb 2003; Sesana et al. 2004; McWilliams et al. 2014; Burke-Spolaor et al. 2019). If all contributing SMBHBs evolve purely by loss of circular orbital energy to gravitational radiation, the resultant GWB spectrum is well described by a simple $f^{-2/3}$ characteristic strain power law (Phinney 2001). However, GWB signals that are not produced by populations of inspiraling black holes may also lie within the nanohertz band; these include primordial GWs from inflation, scalar-induced GWs, and GW signals from multiple processes arising as a result of cosmological phase transitions, such as collisions of bubbles of the post-transition vacuum state, sound waves, turbulence, and the decay of any defects such as cosmic strings or domain walls that may have formed (see, e.g., Guzzetti et al. 2016; Caprini & Figueroa 2018; Domènech 2021, and references therein).

The detection of nanohertz GWs follows the template outlined by Pirani (1956, 2009), whereby we time the propagation of light to measure modulations in the distance between freely falling reference masses. Estabrook & Wahlquist (1975) derived the GW response of electromagnetic signals traveling between Earth and distant spacecraft, sparking interest in low-frequency GW detection. Sazhin (1978) and Detweiler (1979) described nanohertz GW detection using Galactic pulsars and (effectively) the solar system barycenter as references, relying on the regularity of pulsar emission and planetary motions to highlight GW effects. The fact that pulsars are such accurate clocks enables precise measurements of their rotational, astrometric, and binary parameters (and more) from the times of arrival (TOAs) of their pulses, which are used to develop ever-refining end-to-end timing models. Hellings & Downs (1983) made the crucial suggestion that the correlations between the time-of-arrival perturbations of multiple pulsars could reveal a GW signal buried in pulsar noise; Romani (1989) and Foster & Backer (1990) proposed that a pulsar timing array (PTA) of highly stable millisecond pulsars (Backer et al. 1982) could be used to search for a GWB. Nevertheless, the first multipulsar, long-term GWB limits were obtained by analyzing millisecond pulsar residuals independently, rather than as an array (Stinebring et al. 1990; Kaspi et al. 1994).

From a statistical inference standpoint, the problem of detecting nanohertz GWs in PTA data is analogous to GW searches with terrestrial and future space-borne experiments, in which the propagation of light between reference masses is modeled with physical and phenomenological descriptions of signal and noise processes. It is distinguished by the irregular observation times, which encourage a time-domain rather than Fourier-domain formulation, and by noise sources (intrinsic pulsar noise, interstellar-medium-induced radio-frequency-dependent fluctuations, and timing model errors) that are correlated on timescales common to the GWs of interest. This requires the joint estimation of GW signals and noise, which is similar to the kinds of global fitting procedures already used in terrestrial GW experiments and proposed for space-borne experiments. GW analysts have therefore converged on a Bayesian framework that represents all noise sources as Gaussian processes (van Haasteren et al. 2009; van Haasteren & Vallisneri 2014) and relies on model comparison (i.e., Bayes factors, which are ratios of fully marginalized likelihoods) to

define detection (see, e.g., Taylor 2021). This Bayesian approach is nevertheless complemented by null hypothesis testing, using a frequentist detection statistic⁷⁴ (the “optimal statistic” of Anholm et al. 2009; Demorest et al. 2013; Chamberlin et al. 2015) averaged over Bayesian posteriors of the noise parameters (Vigeland et al. 2018).

The GWB—rather than GW signals from individually resolved binary systems—is likely to become the first nanohertz source accessible to PTA observations (Rosado et al. 2015). Because of its stochastic nature, the GWB cannot be identified as a distinctive phase-coherent signal in the way of individual compact-binary-coalescence GWs. Rather, as PTA data sets grow in extent and sensitivity, one expects to first observe the GWB as excess low-frequency residual power of consistent amplitude and spectral shape across multiple pulsars (Pol et al. 2021; Romano et al. 2021). An observation following this behavior was reported in 2020 (Arzoumanian et al. 2020; hereafter NG12gwb) for the 12.5 yr data set collected by the North American Nanohertz Observatory for Gravitational waves (NANOGrav; McLaughlin 2013; Ransom et al. 2019) and then confirmed (Chen et al. 2021; Goncharov et al. 2021b) by the Parkes PTA (PPTA; Manchester et al. 2013) and the European PTA (EPTA; Desvignes et al. 2016), following many years of null results and steadily decreasing upper limits on the GWB amplitude. A combined International PTA (IPTA; Perera et al. 2019) data release consisting of older data sets from the constituent PTAs also confirmed this observation (Antoniadis et al. 2022). Nevertheless, the finding of excess power cannot be attributed to a GWB origin merely by the consistency of amplitude and spectral shape, which could arise from intrinsic pulsar processes of similar magnitude (Goncharov et al. 2022; Zic et al. 2022), or from a common systematic noise such as clock errors (Tiburzi et al. 2016). Instead, definitive proof of GW origin is sought by establishing the presence of phase-coherent interpulsar correlations with the characteristic spatial pattern derived by Hellings and Downs (Hellings & Downs 1983, hereafter HD); for an isotropic GWB, the correlation between the GW-induced timing delays observed at Earth for any pair of pulsars is a universal, quasi-quadrupolar function of their angular separation in the sky. Even though this correlation pattern is modified if there is anisotropy in the GWB—which may be the case for a GWB generated by an SMBHB population (Cornish & Sesana 2013; Mingarelli et al. 2013, 2017; Taylor & Gair 2013; Mingarelli & Sidery 2014; Roebber & Holder 2017)—the HD template is effective for detecting even anisotropic GWBs in all but the most extreme scenarios (Cornish & Sesana 2013; Cornish & Sampson 2016; Taylor et al. 2020; Bécsy et al. 2022; Allen 2023).

In this letter we present multiple lines of evidence for an excess low-frequency signal with HD correlations in the NANOGrav 15 yr data set (Figure 1). Our key results are as follows. The Bayes factor between an HD-correlated, power-law GWB model and a spatially uncorrelated common-spectrum power-law model ranges from 200 to 1000, depending on modeling choices (Figure 2). The noise-marginalized optimal statistic, which is constructed to be selectively sensitive to HD-correlated power, achieves a signal-to-noise ratio (S/N) of ~ 5 (Figures 3 and 4). We calibrated these detection statistics by removing correlations from the 15 yr data set using the phase-

⁷⁴ See Jenet et al. (2006) for an early example of a cross-correlation statistic for PTA GWB detection.

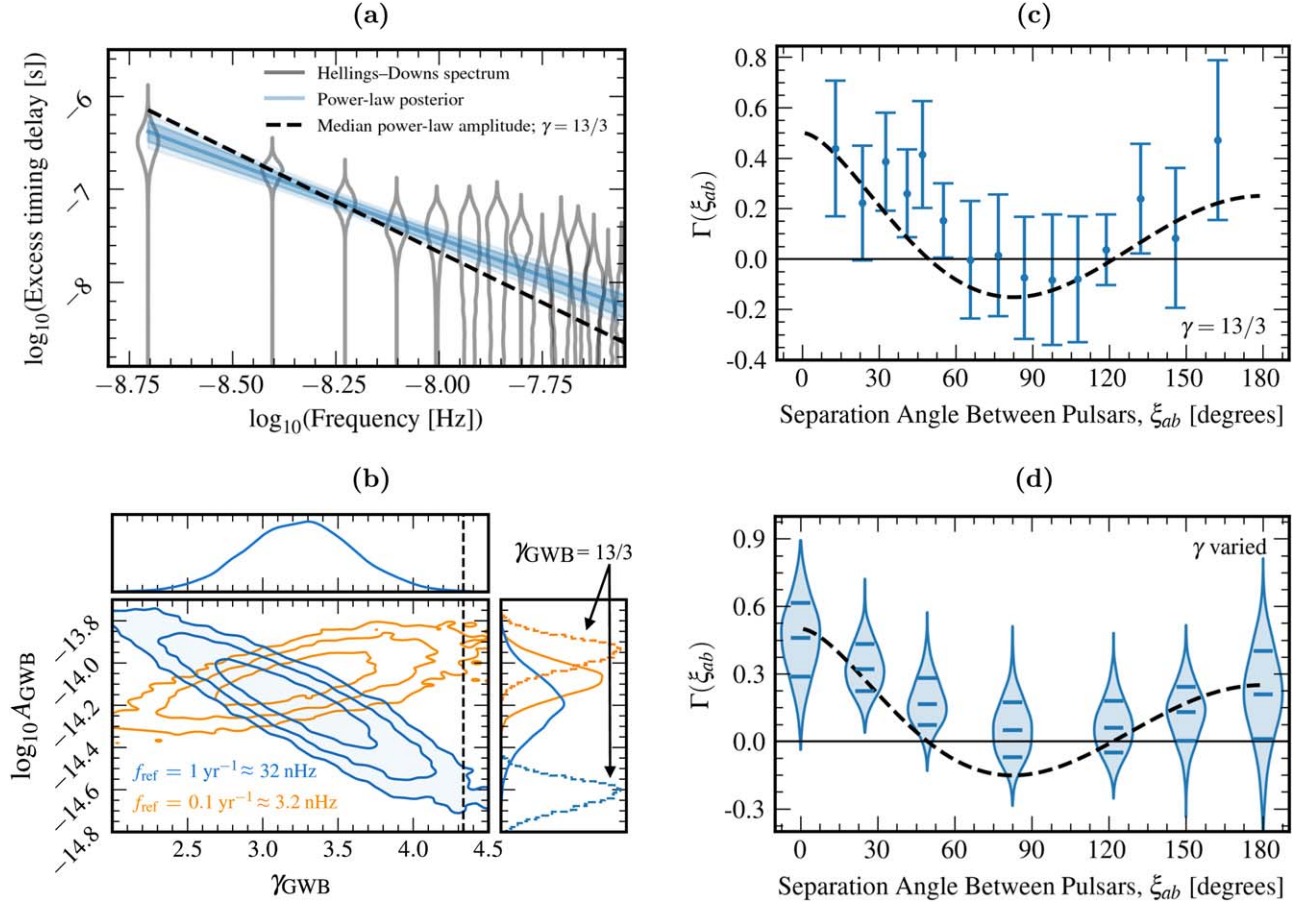


Figure 1. Summary of the main Bayesian and optimal-statistic analyses presented in this paper, which establish multiple lines of evidence for the presence of Hellings–Downs correlations in the 15 yr NANOGrav data set. Throughout we refer to the 68.3%, 95.4%, and 99.7% regions of distributions as $1\sigma/2\sigma/3\sigma$ regions, even in two dimensions. (a) Bayesian “free-spectrum” analysis, showing posteriors (gray violins) of independent variance parameters for a Hellings–Downs-correlated stochastic process at frequencies i/T , with T the total data set time span. The blue represents the posterior median and $1\sigma/2\sigma$ posterior bands for a power-law model; the dashed black line corresponds to a $\gamma = 13/3$ (SMBHB-like) power law, plotted with the median posterior amplitude. See Section 3 for more details. (b) Posterior probability distribution of GWB amplitude and spectral exponent in an HD power-law model, showing $1\sigma/2\sigma/3\sigma$ credible regions. The value $\gamma_{\text{GWB}} = 13/3$ (dashed black line) is included in the 99% credible region. The amplitude is referenced to $f_{\text{ref}} = 1 \text{ yr}^{-1}$ (blue) and 0.1 yr^{-1} (orange). The dashed blue and orange curves in the $\log_{10} A_{\text{GWB}}$ subpanel show its marginal posterior density for a $\gamma = 13/3$ model, with $f_{\text{ref}} = 1 \text{ yr}^{-1}$ and $f_{\text{ref}} = 0.1 \text{ yr}^{-1}$, respectively. See Section 3 for more details. (c) Angular-separation-binned inter-pulsar correlations, measured from 2211 distinct pairings in our 67-pulsar array using the frequentist optimal statistic, assuming maximum-a-posteriori pulsar noise parameters and $\gamma = 13/3$ common-process amplitude from a Bayesian inference analysis. The bin widths are chosen so that each includes approximately the same number of pulsar pairs, and central bin locations avoid zeros of the Hellings–Downs curve. This binned reconstruction accounts for correlations between pulsar pairs (Romano et al. 2021; Allen & Romano 2022). The dashed black line shows the Hellings–Downs correlation pattern, and the binned points are normalized by the amplitude of the $\gamma = 13/3$ common process to be on the same scale. Note that we do not employ binning of inter-pulsar correlations in our detection statistics; this panel serves as a visual consistency check only. See Section 4 for more frequentist results. (d) Bayesian reconstruction of normalized inter-pulsar correlations, modeled as a cubic spline within a variable-exponent power-law model. The violins plot the marginal posterior densities (plus median and 68% credible values) of the correlations at the knots. The knot positions are fixed and are chosen on the basis of features of the Hellings–Downs curve (also shown as a dashed black line for reference): they include the maximum and minimum angular separations, the two zero-crossings of the Hellings–Downs curve, and the position of minimum correlation. See Section 3 for more details.

shift technique, which removes inter-pulsar correlations by adding random phase shifts to the Fourier components of the common process (Taylor et al. 2017). We find false-alarm probabilities of $p = 10^{-3}$ and $p = 5 \times 10^{-5}$ for the observed Bayes factor and optimal statistic, respectively (see Figure 3).

For our fiducial power-law model ($f^{-2/3}$ for characteristic strain and $f^{-13/3}$ for timing residuals) and a log-uniform amplitude prior, the Bayesian posterior of GWB amplitude at the customary reference frequency 1 yr^{-1} is $A_{\text{GWB}} = 2.4^{+0.7}_{-0.6} \times 10^{-15}$ (median with 90% credible interval), which is compatible with current astrophysical estimates for the GWB from SMBHBs (e.g., Burke-Spolaor et al. 2019; Agazie et al. 2023b). This corresponds to a total integrated energy density of $\Omega_{\text{gw}} = 9.3^{+5.8}_{-4.0} \times 10^{-9}$ or $\rho_{\text{gw}} = 7.7^{+4.8}_{-3.3} \times 10^{-17} \text{ ergs cm}^{-3}$ (assuming $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$) in our sensitive frequency

band. For a more general model of the timing-residual power spectral density with variable power-law exponent $-\gamma$, we find $A_{\text{GWB}} = 6.4^{+4.2}_{-2.7} \times 10^{-15}$ and $\gamma = 3.2^{+0.6}_{-0.6}$. See Figure 1(b) for A_{GWB} and γ posteriors. The posterior for γ is consistent with the value of $13/3$ predicted for a population of SMBHBs evolving by GW emission, although smaller values of γ are preferred; however, the recovered posteriors are consistent with predictions from astrophysical models (see Agazie et al. 2023b). We also note that, unlike our detection statistics (which are calibrated under our modeling assumptions), the estimation of γ is very sensitive to minor details in the data model of a few pulsars.

The rest of this paper is organized as follows. We briefly describe our data set and data model in Section 2. Our main results are discussed in detail in Sections 3 and 4; they are supported by a variety of robustness and validation studies,

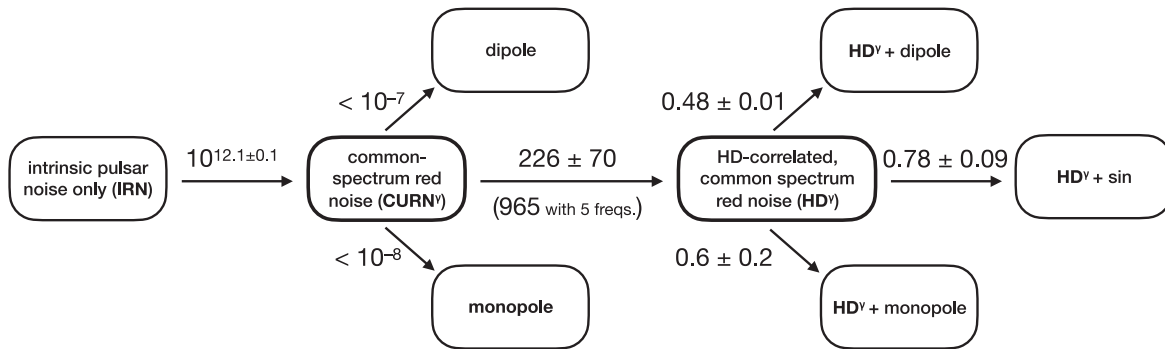


Figure 2. Bayes factors between models of correlated red noise in the NANOGrav 15 yr data set (see Section 5.3 and Appendix B). All models feature variable- γ power laws. CURN^7 is vastly favored over IRN (i.e., we find very strong evidence for common-spectrum excess noise over pulsar intrinsic red noise alone); HD^7 is favored over CURN^7 (i.e., we find evidence for Hellings–Downs correlations in the common-spectrum process); dipole and monopole processes are strongly disfavored with respect to CURN^7 ; adding correlated processes to HD^7 is disfavored. While the interpretation of “raw” Bayes factors is somewhat subjective, they can be given a statistical significance within the hypothesis-testing framework by computing their background distributions and deriving the p -values of the observed factors, e.g., Figure 3.

including a spectral analysis of the excess signal (Section 5.2), a correlation analysis that finds no significant evidence for additional spatially correlated processes (Section 5.3), and cross-validation studies with single-telescope data sets and leave-one-pulsar-out techniques (Section 5.4). In the past 2 yr we have performed an end-to-end review of the NANOGrav experiment, to identify and mitigate possible sources of systematic error or data set contamination; our improvements and considerations are partly described in a set of companion papers: on the NANOGrav statistical analysis as implemented in software (A. Johnson et al. 2023, in preparation), on the 15 yr data set (Agazie et al. 2023a, hereafter NG15), and on pulsar models (Agazie et al. 2023c, hereafter NG15detchar). More companion papers address the possible SMBHB (Agazie et al. 2023b) and cosmological (Afzal et al. 2023) interpretations of our results, with several more GW searches and signal studies in preparation. We look forward to the cross-validation analysis that will become possible with the independent data sets collected by other IPTA members.

2. The 15 yr Data Set and Data Model

The NANOGrav 15 yr data set⁷⁵ (NG15) contains observations of 68 pulsars obtained between 2004 July and 2020 August with the Arecibo Observatory (Arecibo), the Green Bank Telescope (GBT), and the Very Large Array (VLA), augmenting the 12.5 yr data set (Alam et al. 2021a, 2021b) with 2.9 yr of timing data for the 47 pulsars in the previous data set, and with 21 new pulsars.⁷⁶ For this paper we analyze narrowband TOAs, which are computed separately for subbands of each receiver, and focus on the 67 pulsars with a timing baseline ≥ 3 yr. We adopt the TT(BIPM2019) timescale and the JPL DE440 ephemeris (Park et al. 2021), which improves Jupiter’s orbit with ranging and very long baseline interferometry observations of the Juno spacecraft. Uncertainties in the Jovian orbit impacted NANOGrav’s 11 yr GWB search (Arzoumanian et al. 2018; Vallisneri et al. 2020), but they are now negligible.

For each pulsar, we fit the TOAs to a timing model that includes pulsar spin period, spin period derivative, sky location, proper motion, and parallax. While not all pulsars have measurable parallax and proper motion, we always include these parameters because they induce delays with the same frequencies for all pulsars ($f = 0.5 \text{ yr}^{-1}$ for parallax and $f = \text{yr}^{-1}$ plus a linear envelope for proper motion), so there is a risk that a parallax or proper-motion signal could be misidentified as a GW signal. Fitting for these parameters in all pulsars reduces our sensitivity to GWs at those frequencies; however, this effect is minimal for GWB searches since these frequencies are much higher than the frequencies at which we expect the GWB to be significant. For binary pulsars, the timing model includes also five orbital elements for binary pulsars and additional non-Keplerian parameters when these improve the fit as determined by an F -test. We fit variations in dispersion measure (DM) as a piecewise-constant “DMX” function (Arzoumanian et al. 2015; Jones et al. 2017). The individual analysis of each pulsar provides best-fit estimates of the timing residuals δt , of white measurement noise, and of intrinsic red noise, modeled as a power law (Cordes 2013; Jones et al. 2017; Lam et al. 2017).⁷⁷ White measurement noise is described by three parameters: a linear scaling of TOA uncertainties (“EFAC”), white noise added to the TOA uncertainties in quadrature (“EQUAD”), and noise common to all subbands at the same epoch (“ECORR”), with independent parameters for every receiver/back-end combination (see NG15detchar). We summarize white noise by its maximum a posteriori (MAP) covariance C . See Appendix A for more details of our instruments, observations, and data reduction pipeline; a complete discussion of the data set can be found in NG15.

In our Bayesian GWB analysis, we model δt as a finite Gaussian process consisting of time-correlated fluctuations that include intrinsic red pulsar noise and (potentially) a GW signal, along with timing model uncertainties (van Haasteren et al. 2009; van Haasteren & Vallisneri 2014; Taylor 2021). The red noise is modeled with Fourier basis F and amplitudes c (Lentati et al. 2013). All Fourier bases (the columns of F) are sines and cosines computed on the TOAs with frequencies $f_i = i/T$, where $T = 16.03 \text{ yr}$ is the TOA extent. The timing model

⁷⁵ While the time between the first and last observations we analyze is 16.03 yr, this data set is named “15 yr data set” since no single pulsar exceeds 16 yr of observation; we will use this nomenclature despite the discrepancy.

⁷⁶ The data set is available at <http://data.nanograv.org> with the code used to process it.

⁷⁷ Throughout the paper we use “red noise” to describe noise whose power spectrum decreases with increasing frequency.

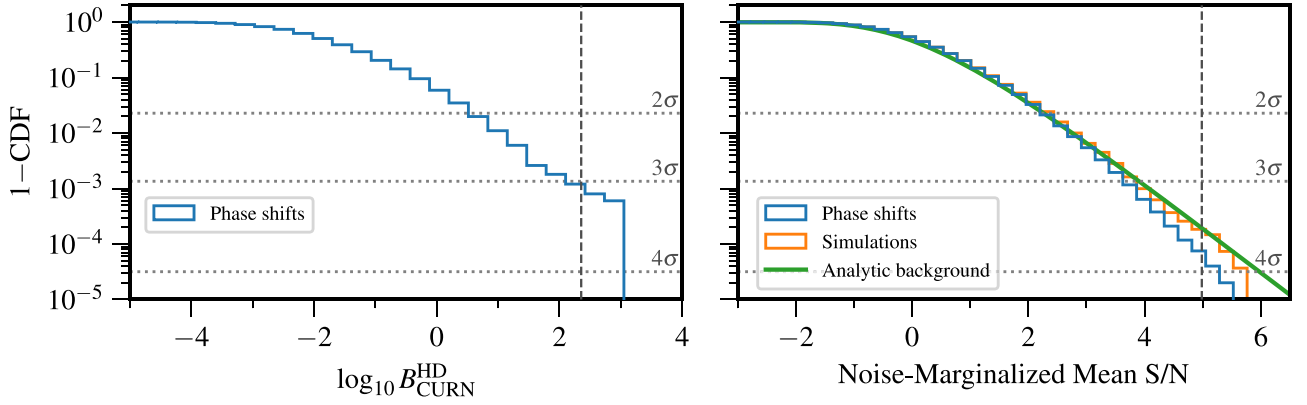


Figure 3. Empirical background distribution of HD^γ -to- CURN^γ Bayes factor (left; see Section 3) and noise-marginalized optimal statistic (right; see Section 4), as computed by the phase-shift technique (Taylor et al. 2017) to remove inter-pulsar correlations. We only compute 5000 Bayesian phase shifts, compared to 400,000 optimal statistic phase shifts, because of the huge computational resources needed to perform the Bayesian analyses. For the optimal statistic, we also compute the background distribution using 27,000 simulations (orange line) and compare to an analytic calculation (green line). Dotted lines indicate Gaussian-equivalent 2σ , 3σ , and 4σ thresholds. The dashed vertical lines indicate the values of the detection statistics for the unshifted data sets. For the Bayesian analyses, we find $p = 10^{-3}$ ($\approx 3\sigma$); for the optimal statistic analyses, we find $p = 5 \times 10^{-5}$ to 1.9×10^{-4} ($\approx 3.5\sigma$ – 4σ).

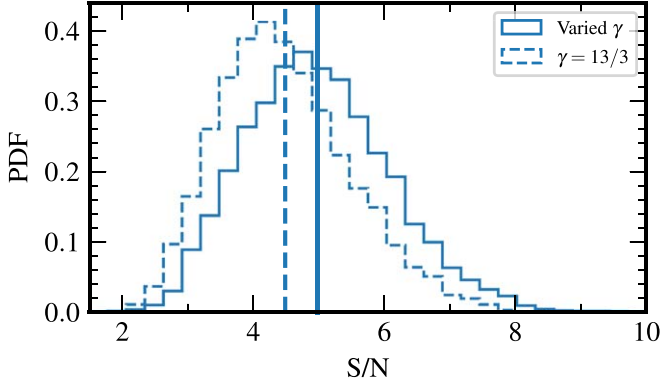


Figure 4. Optimal statistic S/N for HD correlations, distributed over CURN^γ (solid lines) and $\text{CURN}^{13/3}$ (dashed lines) noise parameter posteriors. The vertical lines indicate the mean S/Ns. We find S/Ns of 5 ± 1 and 4 ± 1 for CURN^γ and $\text{CURN}^{13/3}$, respectively.

uncertainties are modeled with design matrix basis $\mathbf{M}$ and coefficients ϵ . The single-pulsar log likelihood is then

$$\ln p(\delta t | \mathbf{c}, \epsilon) = -\frac{1}{2} [\mathbf{r}^T \mathbf{C}^{-1} \mathbf{r} + \ln \det(2\pi \mathbf{C})], \quad (1)$$

with

$$\mathbf{r} = \delta \mathbf{t} - \mathbf{F} \mathbf{c} - \mathbf{M} \epsilon. \quad (2)$$

The prior for the ϵ is taken to be uniform with infinite extent, so the posterior is driven entirely by the likelihood. The set of the $\{\mathbf{c}\}$ for all pulsars take a joint normal prior with zero mean and covariance

$$\langle c_{ai} c_{bj} \rangle = \delta_{ij} (\delta_{ab} \varphi_{ai} + \Phi_{ab,i}); \quad (3)$$

here a, b range over pulsars and i, j over Fourier components, and δ_{ij} is Kronecker's delta. The term φ_{ai} describes the spectrum of intrinsic red noise in pulsar a , while $\Phi_{ab,i}$ describes processes with common spectrum across all pulsars and (potentially) phase-coherent inter-pulsar correlations. The $\{\mathbf{c}\}$ prior ties together the single-pulsar likelihoods given by Equation (1) into a joint posterior, $p(\mathbf{c}, \epsilon, \eta | \delta \mathbf{t}) \propto p(\delta \mathbf{t} | \mathbf{c}, \epsilon) p(\mathbf{c}, \epsilon | \eta) p(\eta)$, where we have dropped subscripts to denote the concatenation of vectors for all pulsars, and where η denotes all

the hyperparameters (such as red-noise and GWB power-spectrum amplitudes) that determine the covariances. We marginalize over $\mathbf{c}$ and ϵ analytically and use Markov chain Monte Carlo (MCMC) techniques (see Appendix B) to estimate $p(\eta | \delta \mathbf{t})$ for different models of the intrinsic red noise and common spectrum.

The data model variants adopted in this paper all share this probabilistic setup but differ in the structure and parameterization of $\Phi_{ab,i}$. For a model with intrinsic red noise only (henceforth IRN), $\Phi_{ab,i} = 0$; for common-spectrum spatially uncorrelated red noise (CURN), $\Phi_{ab,i} = \delta_{ab} \Phi_{\text{CURN},i}$; for an isotropic GWB with Hellings–Downs correlations (HD), $\Phi_{ab,i} = \Gamma(\xi_{ab}) \Phi_{\text{HD},i}$, with Γ the Hellings–Downs function of pulsar angular separations ξ_{ab} ,

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

$$x = \frac{1 - \cos \xi_{ab}}{2}. \quad (5)$$

In NG12gwb we established strong Bayesian evidence for CURN over IRN; finding that HD is preferred over CURN would point to the GWB origin of the common-spectrum signal. We also investigate other spatial correlation patterns, e.g., monopole or dipole, introduced in Section 5.3.

Throughout this paper, we set the spectral components φ_{ai} of intrinsic pulsar noise (which have units of s^2 , as appropriate for the variance of timing residuals) to a power law,

$$\varphi_{ai} = \frac{A_a^2}{12\pi^2} \frac{1}{T} \left(\frac{f_i}{f_{\text{ref}}} \right)^{-\gamma_a} f_{\text{ref}}^{-3}, \quad (6)$$

introducing two dimensionless hyperparameters for each pulsar: the intrinsic noise amplitude A_a and spectral index γ_a . We use log-uniform and uniform priors, respectively, on these hyperparameters; their bounds are described in Appendix B. More sophisticated intrinsic noise models are discussed in Section 5.1 and NG15detchar. In models CURN^γ and HD^γ , the common spectra $\Phi_{\text{CURN},i}$ and $\Phi_{\text{HD},i}$ follow Equation (6),

$$\Phi_{\text{CURN},i} = \frac{A_{\text{CURN}}^2}{12\pi^2} \frac{1}{T} \left(\frac{f_i}{f_{\text{ref}}} \right)^{-\gamma_{\text{CURN}}} f_{\text{ref}}^{-3}, \quad (7)$$

$$\Phi_{\text{HD},i} = \frac{A_{\text{HD}}^2}{12\pi^2} \frac{1}{T} \left(\frac{f_i}{f_{\text{ref}}} \right)^{-\gamma_{\text{HD}}} f_{\text{ref}}^{-3}, \quad (8)$$

introducing hyperparameters A_{CURN} , γ_{CURN} and A_{HD} , γ_{HD} , respectively. However, we set $\gamma_{\text{HD}} = 13/3$ for the GWB from a stationary ensemble of inspiraling binaries and refer to that fiducial model as $\text{HD}^{13/3}$. For specific “free-spectrum” studies we will instead model the individual $\Phi_{\text{CURN},i}$ or $\Phi_{\text{HD},i}$ elements and refer to models $\text{CURN}^{\text{free}}$ and HD^{free} . Throughout this article we use frequencies $f_i = i/T$ with $i = 1-30$ for intrinsic noise ($f = 2-59$ nHz), covering a frequency range over which pulsar noise transitions from red noise dominated to white noise dominated. For common-spectrum noise, we limit the frequency range in order to reduce correlations with excess white noise at higher frequencies. Following NG12gwb, we fit a CURN^γ model enhanced with a power-law break to our data and limit frequencies to the MAP break frequencies ($i = 1-14$ or $f = 2-28$ nHz; see Appendix C).

3. Bayesian Analysis

When fit to the 15 yr data set, the CURN^γ and HD^γ models agree on the presence of a loud time-correlated stochastic signal with common amplitude and spectrum across pulsars.⁷⁸ The joint $A_{\text{HD}}-\gamma_{\text{HD}}$ Bayesian posterior is shown in Figure 1(b), with 1D marginal posteriors in the horizontal and vertical subplots. The posterior medians and 5%–95% quantiles are $A_{\text{HD}} = 6.4_{-2.7}^{+4.2} \times 10^{-15}$ and $\gamma_{\text{HD}} = 3.2_{-0.6}^{+0.6}$. The thicker curve in the vertical subplot is the A_{HD} posterior for the $\text{HD}^{13/3}$ model, for which $A_{\text{HD},13/3} = 2.4_{-0.6}^{+0.7} \times 10^{-15}$. These amplitudes are compatible with astrophysical expectations of a GWB from inspiraling SMBHBs (see Section 6). The A_{HD} posterior has essentially no support below 10^{-15} .

The strong $A_{\text{HD}}-\gamma_{\text{HD}}$ correlation is an artifact of using the $f_{\text{ref}} = 1 \text{ yr}^{-1}$ in Equation (6), and it largely disappears when f_{ref} is moved to the band of greatest PTA sensitivity; see the dashed contours in Figure 1(b) for $f_{\text{ref}} = (10 \text{ yr})^{-1}$. The γ_{HD} posterior is in moderate tension with the theoretical universal binary inspiral value $\gamma_{\text{HD}} = 13/3$, which lies at the 99% credible boundary: smaller values of γ_{HD} could be an indication that astrophysical effects, such as stellar scattering and gas dynamics, play a role in the evolution of SMBHBs emitting GWs in this frequency range (see Section 6; Agazie et al. 2023b). This highlights the importance of measuring this parameter. Furthermore, its estimation is sensitive to details in the modeling of intrinsic red noise and of interstellar medium timing delays in a few pulsars (see the analysis in Section 5.2). Notably, in the 12.5 yr data set $\gamma_{\text{HD}} = 13/3$ was recovered at $\sim 1\sigma$ below the median (NG12gwb); this anomaly is reversed in the newer data set. It is likely that more expansive data sets or more sophisticated chromatic noise models, e.g., next-generation Gaussian process models such as in Section 5.1 (Lam et al. 2018; Goncharov et al. 2021a; Chalumeau et al. 2022), will be needed to infer the presence of possible systematic errors in γ_{HD} .

Our Bayesian analysis provides evidence that the common-spectrum signal includes Hellings–Downs interpulsar correlations. Specifically, the Bayes factor between the HD^γ and CURN^γ models ranges from 200 (when 14 Fourier frequencies

are included in Φ_i) to 1000 (when five frequencies are included, as in NG12gwb). Results are similar for $\text{HD}^{13/3}$ versus $\text{CURN}^{13/3}$. Figure 2 recapitulates Bayes factors between a variety of models, including some with the alternative spatial correlation structures discussed in Section 5.3. The very peaked A_{HD} posterior in Figure 1(b), significantly separated from smaller amplitudes, supports the very large Bayes factor between IRN and CURN^γ . The 15 yr data set favors HD^γ over CURN^γ , as well as over models with monopolar or dipolar correlations, and it is inconclusive about, i.e., gives roughly even odds for, the presence of spatially correlated signals in addition to HD^γ .

We can also regard the HD^γ versus CURN^γ Bayes factor as a detection statistic in a hypothesis-testing framework and derive the p -value of the observed Bayes factor with respect to its empirical distribution under the CURN^γ model. We do so by computing Bayes factors on 5000 bootstrapped data sets where interpulsar spatial correlations are removed by introducing random phase shifts, drawn from a uniform distribution from 0 to 2π , to the common-process Fourier components (Taylor et al. 2017). This procedure alters interpulsar correlations to have a mean of zero, while leaving the amplitudes of intrinsic pulsar noise and CURN unchanged, thus providing a way to test the null hypothesis that no interpulsar correlations are present. The resulting background distribution of Bayes factors is shown in the left panel of Figure 3—they exceed the observed value in 5 of the 5000 phase shifts ($p = 10^{-3}$). We also performed sky scramble analyses (Cornish & Sampson 2016), which remove the dependence of interpulsar spatial correlations on the angular separations between the pulsars by attributing random sky positions to the pulsars. Sky scrambles generate a background distribution for which interpulsar correlations are present in the data, but they are independent of the pulsars’ angular separations; for this distribution, we find $p = 1.6 \times 10^{-3}$. A detailed discussion of sky scrambles and the results of these analyses can be found in Appendix F.

As in NG12gwb, we also carried out a minimally modeled Bayesian reconstruction of the interpulsar correlation pattern, using spline interpolation over seven spline-knot positions. The choice of seven spline-knot positions is based on features of the Hellings–Downs pattern: two correspond to the maximum and minimum angular separations (0° and 180° , respectively), two are chosen to be at the theoretical zero-crossings of the Hellings–Downs pattern (49.2° and 121.8°), one is at the theoretical minimum (82.5°), and the final two are between the end points and zero-crossings (25° and 150°) to allow additional flexibility in the fit. Figure 1(d) shows the marginal 1D posterior densities at these spline-knot positions for a power-law varied-exponent model. The reconstruction is consistent with the overplotted Hellings–Downs pattern; furthermore, the joint 2D marginal posterior densities for the knots, not shown in Figure 1(d), at the HD zero-crossings are consistent with (0, 0) within 1σ credibility.

4. Optimal Statistic Analysis

We complement our Bayesian search with a frequentist analysis using the *optimal statistic* (Anholm et al. 2009; Demorest et al. 2013; Chamberlin et al. 2015), a summary statistic designed to measure correlated excess power in PTA residuals. (Note that there is no accepted definition of “optimal statistic” in modern statistical usage, but the term has become established in the PTA literature to refer to this specific method,

⁷⁸ See Appendix B for details about our Bayesian methods, including the calculation of Bayes factors.

so we use it for this reason.) It is enlightening to describe the optimal statistic as a weighted average of the interpulsar correlation coefficients

$$\rho_{ab} = \frac{\delta \mathbf{t}_a^T \mathbf{P}_a^{-1} \tilde{\Phi}_{ab} \mathbf{P}_b^{-1} \delta \mathbf{t}_b}{\text{Tr } \mathbf{P}_a^{-1} \tilde{\Phi}_{ab} \mathbf{P}_b^{-1} \tilde{\Phi}_{ba}}, \quad (9)$$

where $\delta \mathbf{t}_a^T$ are the residuals of pulsar a and $\mathbf{P}_a = \langle \delta \mathbf{t}_a \delta \mathbf{t}_a^T \rangle$ is their total autocovariance matrix. The cross-covariance matrix $\tilde{\Phi}_{ab}$ encodes the spectrum of the HD-correlated signal, normalized so that $\tilde{\Phi}_{ab} = A^2 \Gamma(\xi_{ab}) \tilde{\Phi}_{ab}$ (see Pol et al. 2022), and where elements of $\tilde{\Phi}_{ab}$ are given by Equation (3). Indeed, the ρ_{ab} have expectation value $A^2 \Gamma(\xi_{ab})$, but their variance $\sigma_{ab}^2 = (\text{Tr } \mathbf{P}_a^{-1} \tilde{\Phi}_{ab} \mathbf{P}_b^{-1} \tilde{\Phi}_{ba})^{-1} + O(A^4)$ is too large to use them directly as estimators. Thus, we assemble the optimal statistic as the variance-weighted, Γ -template-matched average of the ρ_{ab} ,

$$\hat{A}^2 = \frac{\sum_{a>b} \rho_{ab} \Gamma(\xi_{ab}) / \sigma_{ab}^2}{\sum_{a>b} \Gamma^2(\xi_{ab}) / \sigma_{ab}^2}. \quad (10)$$

This equation represents the optimal estimator of the HD amplitude A^2 ; it can also be interpreted as the best-fit $\hat{A}^2$ obtained by least-squares fitting the ρ_{ab} to the Hellings–Downs model $\hat{A}^2 \Gamma(\xi_{ab})$. Because $\hat{A}^2$ is a function of intrinsic red noise and common-process hyperparameters through the $\mathbf{P}_a$, we use the results of an initial Bayesian inference run to refer the statistic to MAP hyperparameters, or to marginalize it over their posteriors. As discussed in Vigeland et al. (2018), we obtain more accurate values of the amplitude by this marginalization.

To search for interpulsar correlations using the optimal statistic, we evaluate the frequency (the p -value) with which an uncorrelated common-spectrum process with parameters estimated from our data set would yield $\hat{A}^2$ greater than we observe. In the absence of a signal, the expectation value of $\hat{A}^2$ is zero, and its distribution is approximately normal. Thus, we divide the observed $\hat{A}^2$ by its standard deviation to define a formal S/N

$$\text{S/N} = \frac{\sum_{a>b} \rho_{ab} \Gamma(\xi_{ab}) / \sigma_{ab}^2}{\left[\sum_{a>b} \Gamma^2(\xi_{ab}) / \sigma_{ab}^2 \right]^{1/2}}. \quad (11)$$

Figure 4 shows the distribution of this S/N over CURN $^\gamma$ and CURN $^{13/3}$ noise parameter posteriors, with S/Ns of 5 ± 1 and 4 ± 1 , respectively (means $\pm$ standard deviations across noise parameter posteriors). We use 14 frequency components to model the signal; the dependence on the number of frequency components is very weak.

Because the distribution of $\hat{A}^2$ is only approximately normal (Hazboun et al. 2023), the S/N of Equation (11) does not map analytically to a p -value, and it cannot be interpreted as a “sigma” level. Instead, optimal statistic p -values can be computed empirically by removing interpulsar correlations from the 15 yr data set with phase shifts (Taylor et al. 2017). We draw random phase offsets from 0 to 2π for the common-process Fourier components, which is equivalent to making uniform draws from the background distribution of CURN, and ask how often a random choice of phase offsets produces an HD-correlated signal. The right panel of Figure 3 shows the distribution of noise-marginalized S/N over 400,000 phase

shifts. There are 19 phase shifts with noise-marginalized S/N greater than observed, with $p = 5 \times 10^{-5}$. We compare the phase-shift distribution with backgrounds obtained by simulation (right panel of Figure 3, orange line) and analytic calculation (green line). For the former, we simulate 27,000 CURN $^\gamma$ realizations using MAP hyperparameters from the 15 yr data and compute the optimal statistic S/N for each; for the latter, we evaluate the generalized χ^2 distribution (Hazboun et al. 2023) with median CURN $^\gamma$ hyperparameters. Although neither method includes the marginalization over noise parameter posteriors, we find good agreement with phase shifts, with $p = 1.8 \times 10^{-4}$ from simulations and $p = 1.9 \times 10^{-4}$ from the analytic calculation. Finally, we use sky scrambles to compute the p -value for the null hypothesis that interpulsar correlations are present, but they have no dependence on the angular separation between the pulsars, for which we find $p < 10^{-4}$ (see Appendix F).

Averaging the cross-correlations ρ_{ab} in angular separation bins with equal numbers of pulsar pairs reveals the Hellings–Downs pattern, as shown in Figure 1(c) for 15 bins. The ρ_{ab} were evaluated with MAP CURN $^{13/3}$ noise parameters. The black dashed curve traces the expected correlations for an HD-correlated background with the MAP amplitude; the vertical error bars display the expected 1σ spreads of the binned cross-correlations, accounting for the $\langle \rho_{ab} \rho_{cd} \rangle$ covariances induced by the HD-correlated process (Romano et al. 2021; Allen & Romano 2022). (Neglecting those covariances yields 20%–40% smaller spreads. Note that they are *not* included in p -value estimates because those are calculated under the null hypothesis of no spatially correlated process.)

Although each draw from the noise parameter posterior would generate a slightly different plot, as would different binnings, the quality of the fit seen in Figure 1 provides a visual indication that the excess low-frequency power in the 15 yr data set harbors HD correlations. The χ^2 for this 15-bin reconstruction with respect to the Hellings–Downs curve is 8.1, where we account for ρ_{ab} covariance in constructing the bins and the covariance between bins in constructing the χ^2 (Allen & Romano 2022). This corresponds to a p -value of 0.75, calculated using simulations based on the HD $^\gamma$ model, or 0.92 if one assumes that this value follows a canonical χ^2 with 15 degrees of freedom. These p -values are representative of what we find with different binnings: we find $p > 0.3$ when using 8–20 bins (assuming a canonical χ^2 distribution).

5. Checks and Validation

Prior to analyzing the 15 yr data set, we extensively reviewed our data collection and analysis procedures, methods, and tools, in an effort to eliminate contamination from systematic effects and human error. Furthermore, the results presented in Sections 3 and 4 are supported by a variety of consistency checks and auxiliary studies. In this section we present those that offer evidence for or against the presence of HD correlations, reveal anomalies, or otherwise highlight features of note in the data: alternative DM modeling (Section 5.1), the spectral content (Section 5.2), and correlation pattern (Section 5.3) of the excess-noise signal, as well as the consistency of our findings across data set “slices,” pulsars, and telescopes (Section 5.4).

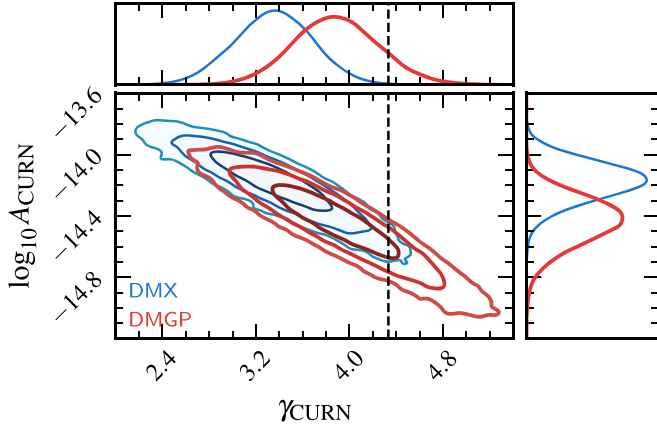


Figure 5. CURN^γ posterior distributions using DMGP (red) and DMX (blue) to model DM variations. The dashed line marks $\gamma_{\text{CURN}} = 13/3$. While the posteriors are broadly consistent, DMGP shifts the γ_{CURN} posterior to higher values, making it more consistent with $\gamma_{\text{CURN}} = 13/3$.

5.1. Alternative DM Models

In this paper and in previous GW searches (e.g., NG12gwb), we model fluctuations in the DM using DMX parameters (a piecewise-constant representation; see NG15). Adopting this DM model as the standard makes it easier to directly compare the results here to those in NG12gwb. An alternative model where DM variations are modeled as a Fourier-domain Gaussian process, DMGP, has been used by Antoniadis et al. (2022), Chen et al. (2021), and Goncharov et al. (2021b). The Fourier coefficients follow a power law similar to those of intrinsic and common-spectrum red noise, but their basis vectors include a ν^{-2} radio frequency dependence, and the component frequencies $f_i = i/T$ range through $i = 1$ –100. Under the DMGP model we also include a deterministic solar-wind model (Hazboun et al. 2022) and the two chromatic events in PSR J1713+0747 reported in Lam et al. (2018), which are modeled as deterministic exponential dips with the chromatic index quantifying the radio frequency dependence of the dips left as a free parameter. If these chromatic events are not modeled, they raise estimated white noise (Hazboun et al. 2020). A detailed discussion of chromatic noise effects can be found in NG15detchar.

Using the DMGP model in place of DMX has minimal effects on nearly all pulsars in the array. Only PSR J1713+0747 and PSR J1600–3053 show notable differences in their recovered intrinsic noise parameters. However, DMGP does affect the parameter estimation of common red noise, as seen in Figure 5, shifting the posterior for γ to higher values that are more consistent with $13/3$. Despite this, we still recover HD correlations at the same significance as when we use DMX to model fluctuations in the DM, implying that the evidence reported for the presence of correlations in this work is independent of the choice of DM noise modeling.

5.2. Spectral Analysis

Adopting power-law spectra for CURN and HD is a useful simplification that reduces the number of fit parameters and yields more informative constraints; furthermore, it is expedient to identify $\text{HD}^{13/3}$ with the hypothesis that we are observing the GWB from SMBHBs. Nevertheless, the standard $\gamma = 13/3$ power law for GW inspirals may be altered by astrophysical processes such as stellar and gas friction in nuclei (see, e.g.,

Merritt & Milosavljević 2005 for a review), by appreciable eccentricity in SMBHB orbits (Enoki & Nagashima 2007), and by low-number SMBHB statistics (Sesana et al. 2008). HD^γ parameter recovery may also be biased if intrinsic pulsar noise is not modeled well by a power law. Indeed, our data show hints of a discrepancy from the idealized $\text{HD}^{13/3}$ model: the γ_{HD} posterior in Figure 1(b) favors slopes much shallower than $13/3$, and the HD^γ -to- CURN^γ Bayes factor drops from 1000 to 200 when Fourier components at more than five frequencies are included in the model.

We examine the spectral content of the 15 yr data set using the $\text{CURN}^{\text{free}}$ and HD^{free} models, which are parameterized by the variances of the Fourier components at each frequency. Their marginal posteriors are shown in the left panel of Figure 6, where bin number i corresponds to $f_i = i/T$, with $T = 16.03$ yr the extent of the data set. For the purpose of illustration, we overlay best-fit power laws that thread the posteriors in a way similar to the factorized PTA likelihood of Taylor et al. (2022) and Lamb et al. (2023).

We deem excess power, either uncorrelated for $\text{CURN}^{\text{free}}$ or correlated for HD^{free} , to be observed in a bin when the support of the posterior is concentrated away from the lowest amplitudes. No power of either kind is observed above f_8 , consistent with the presence of a floor of white measurement noise. Furthermore, no correlated power is observed in bins 6 and 7, where a power-law model would expect a smooth continuation of the trend of bins 1–5 (see the dashed fit of Figure 6); this may explain the drop in the Bayes factor. However, correlated power reappears in bin 8, pushing the fit toward shallower slopes. Indeed, repeating the fit by omitting subsets of the bins suggests that the low recovered γ_{HD} is due mostly to bin 8 and to the lower-than-expected correlated power found in bin 1. Obviously, excluding those bins leads to higher γ_{HD} estimates.

To explore deviations from a pure power law that may arise from statistical fluctuations of the astrophysical background or from unmodeled systematics (perhaps related to the timing model), in Appendix D we relax the normal c_k prior (see Equation (3)) to a multivariate Student’s t -distribution that is more accepting of mild outliers. The resulting estimate of γ_{CURN} peaks at a higher value and is broader than in CURN^γ , with posterior medians and 5%–95% quantiles of $\gamma_{\text{CURN}} = 3.5^{+1.0}_{-1.0}$.

Similarly, spectral turnovers due to interactions between SMBHBs and their environments can result in reduced GWB power at lower frequencies, which might explain the slightly lower correlated power in bin 1. We investigate this hypothesis in Appendix E using the turnover spectrum of Sampson et al. (2015). For this $\text{CURN}^{\text{turnover}}$ model, the 15 yr data favor a spectral bend below 10 nHz (near f_5), but the Bayes factor against the standard HD^γ is inconclusive.

Future data sets with longer time spans and the comparison of our data set with those of other PTAs should help clarify the astrophysical or systematic origin of these possible spectral features.

5.3. Alternative Correlation Patterns

Sources other than GWs can produce interpulsar residual correlations with spatial patterns other than HD. For example, errors in the solar system ephemerides create time-dependent Roemer delays with dipolar correlations (Roebber 2019; Vallisneri et al. 2020), and errors in the correction of telescope time to an inertial timescale (Hobbs et al. 2012, 2020) create an

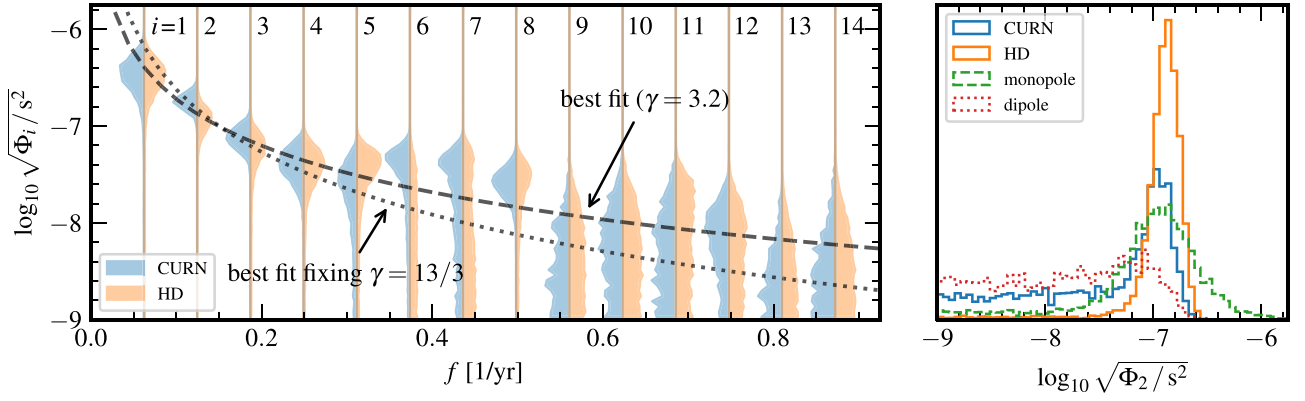


Figure 6. Left: posteriors of Fourier component variance Φ_i for the CURN^{free} (left) and HD^{free} (right) models (see Section 2), plotted at their corresponding frequencies $f_i = i/T$, with T the 16.03 yr extent of the data set. Excess power is observed in bins 1–8 (somewhat marginally in bin 6); Hellings–Downs-correlated power in bins 1–5 and 8. The dashed line plots the best-fit power law, which has $\gamma \simeq 3.2$ (as in Figure 1(d)); the fit is pushed to lower γ by bins 1 and 8. The dotted line plots the best-fit power law when γ is fixed to 13/3; it overshoots in bin 1 and undershoots in bin 8. Right: posteriors of variance Φ_2 in Fourier bin 2 ($f_2 = 3.95$ nHz) in a CURN^{free} + HD^{free} + MONOPOLE^{free} + DIPOLE^{free} model, showing evidence of a quasi-monochromatic monopole process (dashed). No monopole or dipole power is observed in all other bins of that joint model, with $\Phi_{\text{CURN},i}$ and $\Phi_{\text{HD},i}$ posteriors consistent with the left panel.

identical time-dependent delay for all pulsars (i.e., a delay with monopolar correlations).

Gair et al. (2014) showed that, for a pulsar array distributed uniformly across the sky, HD correlations can be decomposed as

$$\Gamma_{\text{HD},ab} = \sum_{l=0}^{\infty} g_l P_l(\cos \xi_{ab}),$$

$$g_0 = 0, \quad g_1 = 0, \quad g_l = \frac{3}{2} (2l+1) \frac{(l-2)!}{(l+2)!} \quad \text{for } l \geq 2, \quad (12)$$

where the $P_l(\cos \xi_{ab})$ are Legendre polynomials of order l evaluated at the pulsar angular separation ξ_{ab} . In other words, an HD-correlated signal should have no power at $l=0$ or $l=1$.

We can perform a frequentist generic correlation search using Legendre polynomials⁷⁹ with the multiple-component optimal statistic (MCOS; Sardesai & Vigeland 2023)—a generalized statistic that allows multiple correlation patterns to be fit simultaneously to the correlation coefficients ρ_{ab} . Figure 7 shows the constraints on $A_l^2 = A^2 g_l$ obtained by fitting the correlations ρ_{ab} to this Legendre series using the MCOS and marginalizing over CURN⁷ noise parameter posteriors. The quadrupolar structure of the data is evident, along with a small but significant monopolar contribution.

The same feature from the Legendre decomposition appears if we use the MCOS to search for multiple correlations simultaneously: a multiple regression analysis favors models that contain both significant HD and monopole correlations (see Appendix G). From simulations of 15 yr–like data sets (see Appendix H.1), we find a p -value of 0.11 ($\sim 2\sigma$) for observing a monopole at this significance or higher with a pure-HD injection of amplitude similar to what we observe. We also perform a model-checking study to assess whether the observed monopole is consistent with the HD^{13/3} model (see Appendix H.2), and we find a p -value of 0.11 for producing an apparent monopole when the signal is purely HD^{13/3}. Thus, we conclude that it is possible for an HD-correlated signal to appear to have monopole correlations in an optimal statistic analysis at this significance level.

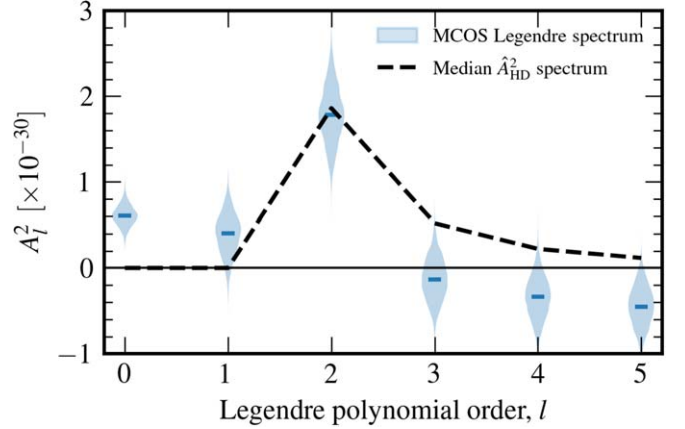


Figure 7. Multiple-component optimal statistic for a Legendre polynomial basis Equation (12) with $l_{\text{max}} = 5$. The violin plots show the distributions of the normalized Legendre coefficients $A_l^2 = A^2 g_l$ over CURN⁷ noise parameter posteriors. The black dashed line shows the Legendre spectrum of a pure-HD signal, with the median posterior $\hat{A}_{\text{HD}}^2$.

In contrast, Bayesian searches for additional correlations do not find evidence of additional monopole- or dipole-correlated red-noise processes; as shown in Figure 2, the Bayes factors for these processes are ~ 1 . We also perform a general Bayesian search for correlations using a CURN^{free} + HD^{free} + MONOPOLE^{free} + DIPOLE^{free} model, which allows for independent uncorrelated and correlated components at every frequency bin. We note that this analysis is more flexible than the ones described above, which assume a power-law power spectral density. We find no significant dipole-correlated power at any frequency, and we find monopole-correlated power only in the second frequency bin ($f_2 = 3.95$ nHz); posteriors of variance for that bin are shown in the right panel of Figure 6.

Motivated by this finding, we perform a search for HD⁷ + SINUSOID, which includes a deterministic sinusoidal delay (applied to all pulsars alike, as appropriate for a monopole) with free frequency, amplitude, and phase. The sinusoid’s posteriors match the free-spectral analysis in frequency and amplitude; however, the Bayes factor between HD⁷ + SINUSOID and HD⁷ calculated using two methods (Hee et al. 2015; Hourihane et al. 2023) is only ~ 1 , so the signal cannot be considered statistically significant. Astrophysically motivated

⁷⁹ A Bayesian method for fitting correlations using Legendre polynomials can be found in Nay et al. (2023).

searches for sources that produce sinusoidal or sinusoid-like delays in the residuals, such as an individual SMBHB or perturbations to the local gravitational field induced by fuzzy dark matter (Khmel'nitsky & Rubakov 2014), also yield Bayes factors ~ 1 . Thus, we conclude that there is some evidence of additional power at 3.95 nHz with monopole correlations; however, the significance in the Bayesian analyses is low, while the optimal statistic S/N could be produced by an HD-correlated signal. Therefore, we cannot definitively say whether the signal is present, or determine the source. We note that performing an MCOS analysis after subtracting off realizations of a sinusoid using $\text{HD}^\gamma + \text{SINUSOID}$ posteriors reduces the $(\text{S/N})_{\text{monopole}} \simeq 0$ while $(\text{S/N})_{\text{HD}}$ remains unchanged, indicating that this single-frequency monopole-correlated signal is likely causing the nonzero monopole signal observed in the MCOS analysis.

Similar hints of a monopolar signal (though weaker) were found in the NANOGrav 12.5 yr data set, unsurprisingly given that it is a subset of the current data set. To exercise due diligence, we audited the correction of telescope time to GPS time at Arecibo and at GBT and found nothing that could explain our observations. The subsequent steps in the time correction pipeline rely on very accurate atomic clocks and are unlikely to introduce considerable systematics (G. Petit 2022, personal communication). An important test will be whether this signal persists in future data sets. If this monopolar feature is truly an astrophysical signal, we would expect it to increase in significance as our data set grows. Comparisons with other PTAs and combined IPTA data sets will also provide crucial insight.

5.4. Dropout and Cross-validation

The GWB is by its nature a signal affecting all of the pulsars in the PTA, although it may appear more significant in some based on their observing time span, noise properties, and the particular realization of pulsar and Earth contributions (Speri et al. 2023). One way to assess the significance of the GWB in each pulsar is a Bayesian *dropout analysis* (Aggarwal et al. 2019; Arzoumanian et al. 2020), which introduces a binary parameter that turns on and off the common signal (or its interpulsar correlations) for a single pulsar, leaving all other pulsars unchanged. The Bayes factor associated with this parameter, also referred to as the “dropout factor,” describes how much each pulsar likes to “participate” in the common signal.

Figure 8 plots CURN^γ versus IRN dropout factors for all 67 pulsars (blue). We find positive dropout factors (i.e., dropout factors > 2) for an uncorrelated common process in 20 pulsars, while only one has a dropout factor < 0.5 . For comparison, in the NANOGrav 12.5 yr data set 10 pulsars showed positive dropout factors for an uncorrelated common process, while three had negative dropout factors. We also show HD correlations versus CURN^γ dropout factors (orange). For these, the uncorrelated common process is always present in all pulsars, but the cross-correlations for all pulsar pairs involving a given pulsar may be dropped from the likelihood. We find positive factors for HD correlations versus CURN^γ in seven pulsars, while three are negative. We expect more pulsars to have positive dropout factors for CURN^γ versus IRN than for Hellings–Downs versus CURN^γ because the Bayes factor comparing the first two models is significantly higher than the one comparing the second two models (see Figure 2).

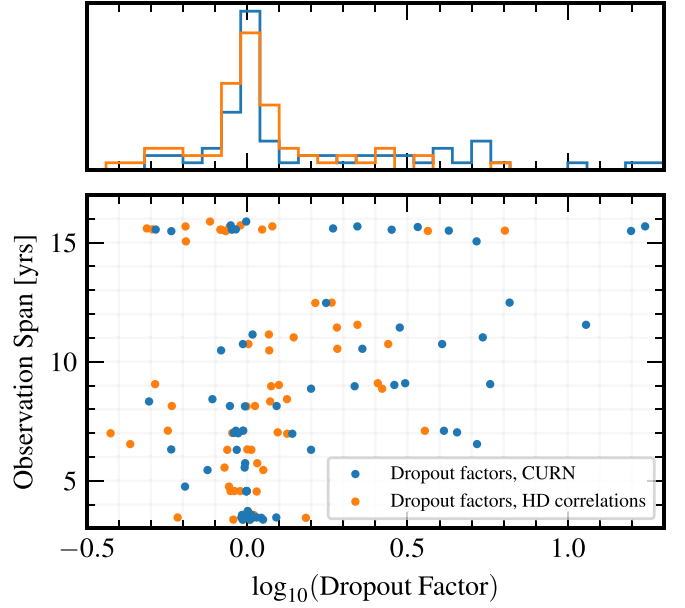


Figure 8. Support for CURN^γ (blue) and HD^γ correlations (orange) in each pulsar, as measured by a dropout analysis. Dropout factors greater than 1 indicate support for the CURN^γ or HD^γ , while those less than 1 show that the pulsar disfavors it. We find significant spread in the dropout factors among pulsars with long observation times, but overall more pulsars favor CURN^γ participation and HD^γ correlations than disfavor them.

Negative dropout factors could be caused by noise fluctuations, or they could be an indication that more advanced chromatic noise modeling is necessary (Alam et al. 2021a). They could also be caused by the GWB itself, which induces both correlated and uncorrelated noise in the pulsars (the so-called “Earth terms” and “pulsar terms”; Mingarelli & Mingarelli 2018).

In addition to Bayes factors, the goodness of fit of probabilistic models can be evaluated by assessing their predictive performance (Gelman et al. 2013). Specifically, given that the GWB is correlated across pulsars, we can (partially) predict the timing residuals δt_a of pulsar a from the residuals δt_{-a} of all other pulsars by way of the “leave-one-out” *posterior predictive likelihood* (PPL)

$$p(\delta t_a | \delta t_{-a}) = \int d\theta_a p(\delta t_a | \theta_a) p(\theta_a | \delta t_{-a}), \quad (13)$$

where θ_a are all the parameters and hyperparameters that affect pulsar a in a given model. As discussed in Meyers et al. (2023), we compare the predictive performance of $\text{CURN}^{13/3}$ and $\text{HD}^{13/3}$ for each pulsar in turn by taking the ratio of the corresponding leave-one-out PPLs. These ratios are closely related to the dropout factors plotted in Figure 8. Multiplying the PPL ratios for all pulsars yields the pseudo-Bayes factor (PBF). For the 15 yr data set we find $\text{PBF}_{15 \text{ yr}} = 1400$ in favor of $\text{HD}^{13/3}$ over $\text{CURN}^{13/3}$. The PBF does not have a “betting odds” interpretation, but we obtain a crude estimate of its significance by building its background distribution on 40 $\text{CURN}^{13/3}$ simulations with the $\text{MAP } \log_{10} A_{\text{CURN}}$ inferred from the 15 yr data set. For all simulations except one, the PBF favors the null hypothesis, and $\log_{10} \text{PBF}_{15 \text{ yr}}$ is displaced by approximately three standard deviations from the mean $\log_{10} \text{PBF}$.

A different sort of cross-validation relies on evaluating the optimal statistic for temporal subsets of the data set, as in

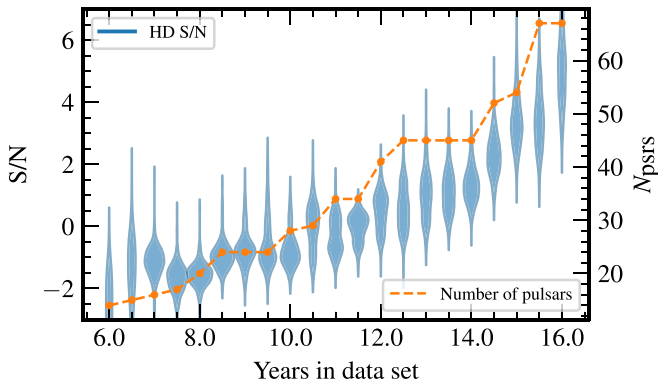


Figure 9. S/N growth as a function of time and number of pulsars. As we move from left to right we add an additional 6 months of data at each step. New pulsars are added when they accumulate 3 yr of data. The blue violin plot shows the distribution of the optimal statistic S/N over CURN^γ noise parameters. The dashed orange line shows the number of pulsars used for each time slice.

Hazboun et al. (2020). In the regime where the lowest frequencies of our data are dominated by the GWB, the optimal statistic S/N should grow with the square root of the time span of the data and linearly with the number of pulsars in the array (Siemens et al. 2013); in this regime increasing the number of pulsars is the best way to boost PTA sensitivity to the GWB. To verify that this is indeed the case, we analyze “slices” of the data set in 6-month increments, starting from a 6 yr data set. Once a new pulsar accumulates 3 yr of data, we add it to the array. We perform a separate Bayesian CURN^γ analysis for each slice and calculate the Hellings–Downs optimal statistic over the noise parameter posterior. In Figure 9, we plot the S/N distributions against time span and the number of pulsars. As expected, we observe essentially monotonic growth associated with the increase in the number of pulsars.

The signal should also be consistent between timing observations made with Arecibo and GBT. To test this, we analyze the two split-telescope data sets (see Appendix A); both show evidence of common-spectrum excess noise. Figure 10 shows Arecibo (orange) and GBT (green) CURN^γ posteriors, which are broadly consistent with each other and with full-data posteriors (blue). Arecibo yields $\log_{10} A = -14.02^{+0.18}_{-0.22}$ and $\gamma = 2.78^{+0.70}_{-0.64}$ (medians with 68% credible intervals), while GBT yields $\log_{10} A = -14.2^{+0.15}_{-0.17}$ and $\gamma = 3.37^{+0.40}_{-0.38}$.

The split-telescope data sets are significantly less sensitive to spatial correlations than the full data set because they have fewer pulsars and therefore fewer pulsar pairs. Nevertheless, we can search them for spatial correlations using the optimal statistic. We find a noise-marginalized Hellings–Downs S/N of 2.9 for Arecibo and 3.3 for GBT, consistent with the split-telescope data sets having about half the number of pulsars as the full data set. The S/Ns for Arecibo and GBT are comparable; while telescope sensitivity, observing cadence, and distribution of pulsars all affect GWB sensitivity, the dominant factor is the number of pulsars because the S/N scales linearly with the number of pulsars but only as $\propto(\sigma\sqrt{c})^{-1/\gamma}$, where σ is the residual rms and c is the observing cadence (Siemens et al. 2013). We also note that the distributions of angular separations probed by Arecibo and GBT are similar, although GBT observes more pulsar pairs with large angular separations (see Appendix A).

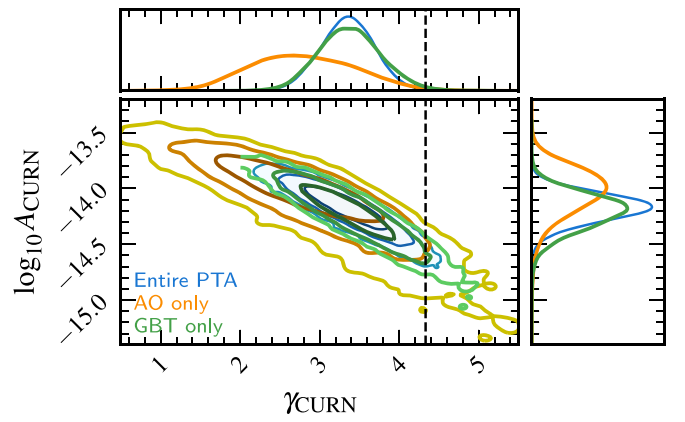


Figure 10. CURN^γ posterior distributions for Arecibo (orange) and GBT (green) split-telescope data sets, and for the full data set (blue). The dashed line marks $\gamma_{\text{CURN}} = 13/3$. The posteriors for the split-telescope data sets are consistent with each other and with the posteriors for the full data set.

6. Discussion

In this letter we have reported on a search for an isotropic stochastic GWB in the 15 yr NANOGrav data set. A previous analysis of the 12.5 yr NANOGrav data set found strong evidence for excess low-frequency noise with common spectral properties across the array but inconclusive evidence for Hellings–Downs interpulsar correlations, which would point to the GW origin of the background. By contrast, the 12.5 yr data disfavored purely monopolar (clock-error-like) and dipolar (ephemeris-error-like) correlations. Subsequent independent analyses by the PPTA and EPTA collaborations reported results consistent with ours (Chen et al. 2021; Goncharov et al. 2021b), as did the search of a combined data set (Antoniadis et al. 2022)—a syzygy of tantalizing discoveries that portend the rise of low-frequency GW astronomy.

We analyzed timing data for 67 pulsars in the 15 yr data set (those that span >3 yr), with a total time span of 16.03 yr, and more than twice the pulsar pairs than in the 12.5 yr data set. The common-spectrum stochastic signal gains even greater significance and is detected in a larger number of pulsars. For the first time, we find compelling evidence of Hellings–Downs interpulsar correlations, using both Bayesian and frequentist detection statistics (see Figure 1), with false-alarm probabilities of $p = 10^{-3}$ and $p = 5 \times 10^{-5}$ to 1.9×10^{-4} , respectively (see Figure 3).

The significance of Hellings–Downs correlations increases as we increase the number of frequency components in the analysis up to five, indicating that the correlated signal extends over a range of frequencies. A detailed spectral analysis supports a power-law signal, but at least two frequency bins show deviations that may skew the determination of spectral slope (Figure 6). These discrepancies may arise from astrophysical or systematic effects. Furthermore, slope determination changes significantly using an alternative DM model (Figure 5). The study of spatial correlations with the optimal statistic confirms a Hellings–Downs quasi-quadrupolar pattern (Figure 7 and Figure 1(c)), with some indications of an additional monopolar signal confined to a narrow frequency range near 4 nHz. However, the Bayesian evidence for this monopolar signal is inconclusive, and we could not ascribe it to any astrophysical or terrestrial source (e.g., an individual SMBHB or errors in the chain of timing corrections).

The GWB is a persistent signal that should increase in significance with number of pulsars and observing time span. This is indeed what we observe by analyzing slices of the data set (see Figure 9). Furthermore, the signal is present in multiple pulsars (Figure 8) and can be found in independent single-telescope data sets (Figure 10). We are preparing a number of other papers searching the 15 yr data set for stochastic and deterministic signals, including an all-sky, all-frequency search for GWs from individual circular SMBHBs. This search, together with the same analysis of the 12.5 yr data set (Arzoumanian et al. 2023), indicates that the spectrum and correlations we observe cannot be produced by an individual circular SMBHB.

If the Hellings–Downs-correlated signal is indeed an astrophysical GWB, its origin remains indeterminate. Among the many possible sources in the PTA frequency band, numerous studies have focused on the unresolved background from a population of close-separation SMBHBs. The SMBHB population is a direct by-product of hierarchical structure formation, which is driven by galaxy mergers (e.g., Blumenthal et al. 1984). In a post-merger galaxy, the SMBHBs sink to the center of the common merger remnant through dynamical interactions with their astrophysical environment, eventually leading to the formation of a binary (Begelman et al. 1980). GW emission from an SMBHB at nanohertz frequencies is quasi-monochromatic because the binaries evolve very slowly. Under the assumption of purely GW-driven binary evolution, the expected characteristic strain spectrum is $\propto f^{-2/3}$ (or $f^{-13/3}$ for pulsar timing residuals).

The GWB spectrum may also feature a low-frequency turnover induced by the dynamical interactions of binaries with their astrophysical environment (e.g., with stars or gas; see Armitage & Natarajan 2002; Sesana et al. 2004; Merritt & Milosavljević 2005), or possibly by nonnegligible orbital eccentricities persisting to small separations (Enoki & Nagashima 2007). We find little support for a low-frequency turnover in our data (see Appendix E).

The GWB amplitude is determined primarily by SMBHB masses and by the occurrence rate of close binaries, which in turn depends on the galaxy merger rate, the occupation fraction of SMBHBs, and the binary evolution timescale; population models predict amplitudes ranging over more than an order of magnitude (Rajagopal & Romani 1995; Jaffe & Backer 2003; Wyithe & Loeb 2003; Sesana 2013; McWilliams et al. 2014), under a variety of assumptions. Figure 11 displays a comparison of HD^γ parameter posteriors with power-law spectral fits from an observationally constrained semianalytic model of the SMBHB population constructed with the HOLODECK package (L. Z. Kelley et al. 2023, in preparation). This particular set of SMBHB populations assumes purely GW-driven binary evolution and uses relatively narrow distributions of model parameters based on literature constraints from galaxy merger observations (see, e.g., Tomczak et al. 2014). While the amplitude recovered in our analysis is consistent with models derived directly from our understanding of SMBHB and galaxy evolution, it is toward the upper end of predictions implying a combination of relatively high SMBHB masses and binary fractions. A detailed discussion of the GWB from SMBHBs in light of our results is given in Agazie et al. (2023b).

In addition to SMBHBs, more exotic cosmological sources such as inflation, cosmic strings, phase transitions, domain

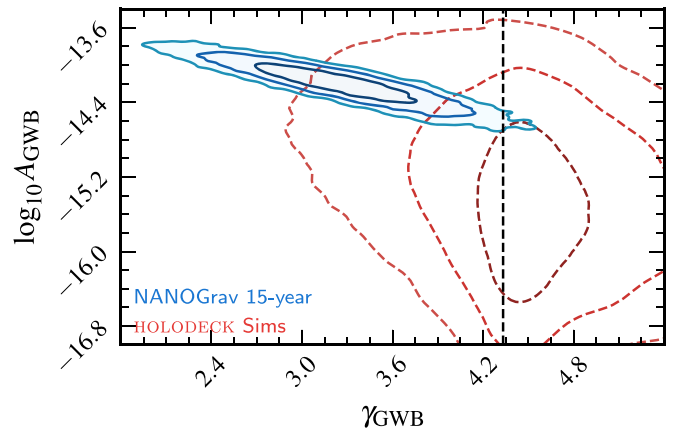


Figure 11. Posteriors of HD^γ amplitude (for $f_{\text{ref}} = 1 \text{ yr}^{-1}$) and spectral slope for the 15 yr data set (blue), compared to power-law fits to simulated GWB spectra (red dashed) from a population of SMBHBs generated by HOLODECK (L. Z. Kelley et al. 2023, in preparation) under the assumption of purely GW-driven binary evolution and narrowly distributed model parameters based on galaxy merger observations. We show $1\sigma/2\sigma/3\sigma$ regions, and the dashed line indicates $\gamma = 13/3$. The broad contours confirm that population variance can lead to a significant spread of spectral characteristics.

walls, and curvature-induced GWs can also produce detectable GWBs in the nanohertz range (see, e.g., Guzzetti et al. 2016; Caprini & Figueroa 2018, and references therein). Similarities in the spectral shapes of cosmological and astrophysical signals make it challenging to determine the origin of the background from its spectral characterization (Kaiser et al. 2022). The question could be settled by the detection of signals from individual loud SMBHBs or by the observation of spatial anisotropies, since the anisotropies expected from SMBHBs are orders of magnitude larger than those produced by most cosmological sources (Caprini & Figueroa 2018; Bartolo et al. 2022). We discuss these models in the context of our results in Afzal et al. (2023).

The EPTA and Indian PTA (InPTA; Joshi et al. 2018), PPTA, and Chinese PTA (CPTA; Lee 2016) Collaborations have also recently searched their most recent data for signatures of a GWB (J. Antoniadis et al. 2023, in preparation; D. J. Reardon et al. 2023, in preparation; H. Xu et al. 2023, in preparation), and an upcoming IPTA paper will compare the results of these searches. The IPTA’s forthcoming Data Release 3 will combine the NANOGrav 15 yr data set with observations from the EPTA, PPTA, and InPTA collaborations, comprising about 80 pulsars with time spans up to 24 yr and offering significantly greater sensitivity to spatial correlations and spectral characteristics than single-PTA data sets. Future PTA observation campaigns will improve our understanding of this signal and of its astrophysical and cosmological interpretation. Longer data sets will tighten spectral constraints on the GWB, clarifying its origin (Pol et al. 2021). Greater numbers of pulsars will allow us to probe anisotropy in the GWB (Pol et al. 2022) and its polarization structure (see, e.g., Arzoumanian et al. 2021, and references therein). The observation of a stochastic signal with spatial correlations in PTA data, suggesting a GWB origin, expands the horizon of GW astronomy with a new galaxy-scale observatory sensitive to the most massive black hole systems in the universe and to exotic cosmological processes.

Acknowledgments

The NANOGrav Collaboration receives support from National Science Foundation (NSF) Physics Frontiers Center award Nos. 1430284 and 2020265, the Gordon and Betty Moore Foundation, NSF AccelNet award No. 2114721, an NSERC Discovery Grant, and CIFAR. The Arecibo Observatory is a facility of the NSF operated under cooperative agreement (AST-1744119) by the University of Central Florida (UCF) in alliance with Universidad Ana G. Méndez (UAGM) and Yang Enterprises (YEI), Inc. The Green Bank Observatory is a facility of the NSF operated under cooperative agreement by Associated Universities, Inc. The National Radio Astronomy Observatory is a facility of the NSF operated under cooperative agreement by Associated Universities, Inc. This work used the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant No. ACI-1548562. Specifically, it used the Bridges-2 system, which is supported by NSF award No. ACI-1928147, at the Pittsburgh Supercomputing Center (PSC). This work was conducted using the Thorny Flat HPC Cluster at West Virginia University (WVU), which is funded in part by National Science Foundation (NSF) Major Research Instrumentation Program (MRI) award No. 1726534 and West Virginia University. This work was also conducted in part using the resources of the Advanced Computing Center for Research and Education (ACCRE) at Vanderbilt University, Nashville, TN. This work was facilitated through the use of advanced computational, storage, and networking infrastructure provided by the Hyak supercomputer system at the University of Washington. This research was supported in part through computational resources and services provided by Advanced Research Computing at the University of Michigan, Ann Arbor. NANOGrav is part of the International Pulsar Timing Array (IPTA); we would like to thank our IPTA colleagues for their feedback on this paper. We thank members of the IPTA Detection Committee for developing the IPTA Detection Checklist. We thank Bruce Allen for useful feedback. We thank Valentina Di Marco and Eric Thrane for uncovering a bug in the sky scramble code. We thank Jolien Creighton and Leo Stein for helpful conversations about background estimation. L.B. acknowledges support from the National Science Foundation under award AST-1909933 and from the Research Corporation for Science Advancement under Cottrell Scholar Award No. 27553. P.R.B. is supported by the Science and Technology Facilities Council, grant No. ST/W000946/1. S. B. gratefully acknowledges the support of a Sloan Fellowship and the support of NSF under award No. 1815664. The work of R.B., R.C., D.D., N.La., X.S., J.P.S., and J.T. is partly supported by the George and Hannah Bolinger Memorial Fund in the College of Science at Oregon State University. M.C., P. P., and S.R.T. acknowledge support from NSF AST-2007993. M.C. and N.S.P. were supported by the Vanderbilt Initiative in Data Intensive Astrophysics (VIDA) Fellowship. K.Ch., A.D. J., and M.V. acknowledge support from the Caltech and Jet Propulsion Laboratory President’s and Director’s Research and Development Fund. K.Ch. and A.D.J. acknowledge support from the Sloan Foundation. Support for this work was provided by the NSF through the Grote Reber Fellowship Program administered by Associated Universities, Inc./National Radio Astronomy Observatory. Support for H.T.C. is provided by NASA through the NASA Hubble Fellowship Program grant No. HST-HF2-51453.001 awarded by the Space Telescope

Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. K.Cr. is supported by a UBC Four Year Fellowship (6456). M.E.D. acknowledges support from the Naval Research Laboratory by NASA under contract S-15633Y. T.D. and M.T.L. are supported by an NSF Astronomy and Astrophysics grant (AAG) award No. 2009468. E.C.F. is supported by NASA under award No. 80GSFC21M0002. G.E.F., S.C.S., and S.J.V. are supported by NSF award PHY-2011772. K.A.G. and S.R.T. acknowledge support from an NSF CAREER award No. 2146016. The Flatiron Institute is supported by the Simons Foundation. S.H. is supported by the National Science Foundation Graduate Research Fellowship under grant No. DGE-1745301. N.La. acknowledges support from the Larry W. Martin and Joyce B. O’Neill Endowed Fellowship in the College of Science at Oregon State University. Part of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004). D.R.L. and M.A.Mc. are supported by NSF No. 1458952. M.A.Mc. is supported by NSF No. 2009425. C.M.F.M. was supported in part by the National Science Foundation under grant Nos. NSF PHY-1748958 and AST-2106552. A.Mi. is supported by the Deutsche Forschungsgemeinschaft under Germany’s Excellence Strategy—EXC 2121 Quantum Universe—390833306. P.N. acknowledges support from the BHL, funded by grants from the John Templeton Foundation and the Gordon and Betty Moore Foundation. The Dunlap Institute is funded by an endowment established by the David Dunlap family and the University of Toronto. K.D.O. was supported in part by NSF grant No. 2207267. T.T.P. acknowledges support from the Extragalactic Astrophysics Research Group at Eötvös Loránd University, funded by the Eötvös Loránd Research Network (ELKH), which was used during the development of this research. S.M. R. and I.H.S. are CIFAR Fellows. Portions of this work performed at NRL were supported by ONR 6.1 basic research funding. J.D.R. also acknowledges support from start-up funds from Texas Tech University. J.S. is supported by an NSF Astronomy and Astrophysics Postdoctoral Fellowship under award AST-2202388 and acknowledges previous support by the NSF under award 1847938. C.U. acknowledges support from BGU (Kreitman fellowship) and the Council for Higher Education and Israel Academy of Sciences and Humanities (Excellence fellowship). C.A.W. acknowledges support from CIERA, the Adler Planetarium, and the Brinson Foundation through a CIERA-Adler postdoctoral fellowship. O.Y. is supported by the National Science Foundation Graduate Research Fellowship under grant No. DGE-2139292.

Dedication. This paper is dedicated to the memory of Donald Backer: a pioneer in PTAs, a term he coined; a discoverer of the first millisecond pulsar; a master developer of pulsar timing instrumentation; a founding member of NANOGrav; and a friend and mentor to many of us.

Facilities: Arecibo, GBT, VLA.

Software: `acor`, `astropy` (Astropy Collaboration et al. 2022), `ceffyl` (Lamb et al. 2023), `chainconsumer` (Hinton 2016), `ENTERPRISE` (Ellis et al. 2019), `enterprise_extensions` (Taylor et al. 2018), `hasasia` (Hazboun et al. 2019), `holodeck` (L. Z. Kelley et al. 2023, in preparation), `Jupyter` (Kluyver et al. 2016), `libstempo` (Vallisneri 2020), `matplotlib` (Hunter 2007), `numpy`

(Harris et al. 2020), PINT (Luo et al. 2021), PTMCMC (Ellis & van Haasteren 2017), `scipy` (Virtanen et al. 2020), `Tempo2` (Hobbs & Edwards 2012).

Author contributions

An alphabetical-order author list was used for this paper in recognition of the fact that a large, decade-timescale project such as NANOGrav is necessarily the result of the work of many people. All authors contributed to the activities of the NANOGrav Collaboration leading to the work presented here and reviewed the manuscript, text, and figures prior to the paper’s submission. Additional specific contributions to this paper are as follows. G.A., A.A., A.M.A., Z.A., P.T.B., P.R.B., H.T.C., K.C., M.E.D., P.B.D., T.D., E.C.F., W.F., E.F., G.E.F., N.G., P.A.G., J.G., D.C.G., J.S.H., R.J.J., M.L.J., D.L.K., M. K., M.T.L., D.R.L., J.L., R.S.L., A.M., M.A.M., N.M., B.W. M., C.N., D.J.N., T.T.P., B.B.P.P., N.S.P., H.A.R., S.M.R., P. S.R., A.S., C.S., B.J.S., I.H.S., K.S., A.S., J.K.S., and H.M. W. developed the 15 yr data set through a combination of observations, arrival time calculations, data checks and refinements, and timing model development and analysis; additional specific contributions to the data set are summarized in NG15. S.R.T. and S.J.V. led the search and coordinated the writing of this paper. P.T.B., J.S.H., P.M.M., N.S.P., J.S., S.R. T., M.V., and S.J.V. wrote the paper, made the figures, and generated the bibliography. K.P.I., N.L., N.S.P., X.S., J.S., J.P. S., S.R.T., and S.J.V. performed analyses and developed new techniques on a preliminary 14 yr data set. P.T.B., B.B., B.D. C., S.C., B.D., G.E.F., K.G., A.D.J., A.R.K., L.Z.K., N.L., W. G.L., N.S.P., S.C.S., L.S., X.S., J.S., J.P.S., S.R.T., C.U., M.V., S.J.V., Q.W., and C.A.W. analyzed preliminary versions of the 15 yr data set. J.S.H., J.G., N.L., N.S.P., J.P.S., and J.K. S. performed noise analyses on the data set. P.T.B., B.B., R.B., R.C., M.C., N.J.C., D.D., H.D., G.E.F., K.A.G., S.H., A.D.J., N.L., W.G.L., P.M.M., P.P., N.S.P., S.C.S., X.S., J.S., J.P.S., J. T., S.R.T., M.V., S.J.V., and C.A.W. performed the Bayesian and frequentist analyses presented in this paper. K.G., J.S.H., P. M.M., J.D.R., and S.J.V. computed the background distribution of our detection statistics. P.R.B., P.M.M., N.S.P., and M. V. performed simulations that were used to compute false-alarm probabilities. A.M.A., P.T.B., B.B., B.D., S.C., J.A.E., G.E.F., J.S.H., A.D.J., A.R.K., N.L., M.T.L., K.D.O., T.T.P., N.S.P., J.S., J.P.S., J.K.S., S.R.T., M.V., and S.J.V. contributed to the development of our code. P.T.B., N.J.C., J.S.H., A.D.J., T.B.L., P.M.M., J.D.R., and M.V. performed an internal code review. K.C., C.C., J.A.E., J.S.H., D.R.M., M.A.Mc., C.M.F. M., K.D.O., N.S.P., S.R.T., M.V., R.v.H., and S.J.V. were part of a NANOGrav Detection Study Group. M.A.M. and P. N. served on the IPTA Detection Committee. K.C., J.S.H., T.J. W.L., P.M.M., N.S.P., J.D.R., X.S., S.R.T., S.J.V., and C.A. W. responded to the Detection Checklist from the IPTA Detection Committee.

Appendix A Additional Data Set Details

The observations included in the NANOGrav 15 yr data set were performed between 2004 July and 2020 August with the 305 m Arecibo Observatory (Arecibo), the 100 m Green Bank Telescope (GBT), and, since 2015, the 27 25 m antennae of the Very Large Array (VLA). We used Arecibo to observe the 33 pulsars that lie within its decl. range ($0^\circ < \delta < +39^\circ$); GBT to

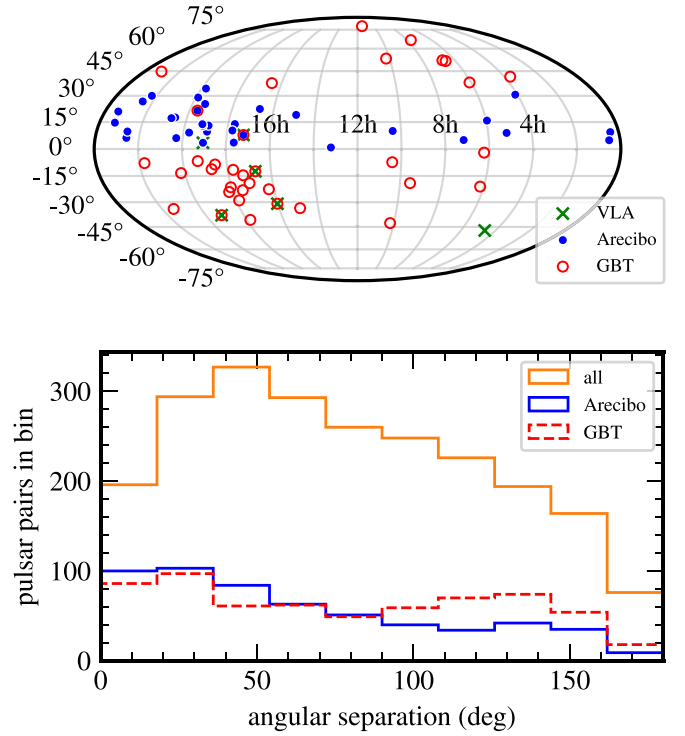


Figure 12. Top: sky locations of the 67 pulsars used in the 15 yr GWB analysis. Markers indicate which telescopes observed the pulsar. Bottom: distribution of angular separations probed by the pulsars in the full data set (orange), the Arecibo data set (blue), and the GBT data set (red). Because Arecibo and GBT mostly observed pulsars at different declinations, there are few intertelescope pairs at small angular separations, resulting in a deficit of pairs for the full data set in the first bin.

observe the pulsars that lie outside of Arecibo’s range, plus J1713 + 0747 and B1937 + 21, for a total of 36 pulsars; and the VLA to observe the seven pulsars J0437–4715, J1600–3053, J1643–1224, J1713 + 0747, J1903 + 0327, J1909–3744, and B1937 + 21. Six of these were also observed with Arecibo, GBT, or both; J0437–4715 was only visible to the VLA. Figure 12 shows the sky locations of the 67 pulsars used for the GWB search (top) and the distribution of angular separations for the pulsar pairs (bottom).

Initial observations were performed with the ASP (Arecibo) and GASP (GBT) systems, with 64 MHz bandwidth (Demorest 2007). Between 2010 and 2012, we transitioned to the PUPPI (Arecibo) and GUPPI (GBT) systems, with bandwidths up to 800 MHz (DuPlain et al. 2008; Ford et al. 2010). We observe pulsars in two different radio frequency bands in order to measure pulse dispersion from the interstellar medium: at Arecibo, we use the 1.4 GHz receiver plus either the 430 MHz or 2.1 GHz receiver (and the 327 MHz receiver for early observations of J2317+1439); at GBT, we use the 820 MHz and 1.4 GHz receivers; at the VLA, we use the 1.4 and 3 GHz receivers with the YUPPI system.

In Section 5.4 we also analyze two split-telescope data sets: 33 pulsars for Arecibo, and 35 for GBT (excluding J0614–3329, which was observed for less than 3 yr). For the two pulsars timed by both telescopes (J1713+0747 and B1937+21), we partition the timing data between the telescopes and obtain independent timing solutions for each. We do not analyze a VLA-only data set, which would have shorter observation spans and significantly reduced sensitivity.

Appendix B

Bayesian Methods and Diagnostics

The prior probability distributions assumed for all analyses in this paper are listed in Table 1. We use MCMC techniques to sample randomly from the joint posterior distribution of our model parameters. Marginal distributions are obtained simply by considering only the parameter of interest in each sample. To assess convergence of our MCMC runs beyond visual inspection, we use the Gelman–Rubin statistic, requiring $\hat{R} < 1.01$ for all parameters (Gelman & Rubin 1992; Vehtari et al. 2021). We performed most runs discussed in this paper with the PTMCMC sampler (Ellis & van Haasteren 2017) and post-processed samples with `chainconsumer` (Hinton 2016).

In NG12gwb we use an analytic approximation for the uncertainty of marginalized posterior statistics (Wilcox 2012). Here we instead adopt a bootstrap approach: we resample the original MCMC samples (with replacement) to generate new sets that act as independent sampling realizations. We then calculate the distributions of the desired summary statistics (e.g., quantiles and marginalized posterior values) over these sets. From these distributions, we determine central values and uncertainties (either medians and 68% confidence intervals, or means and standard deviations).

We rely on a variety of techniques to perform Bayesian model comparison. The first is thermodynamic integration (e.g., Ogata 1989; Gelman & Meng 1998), which computes Bayesian evidence integrals directly through parallel tempering: we run N_β MCMC chains that explore variants of the likelihood function raised to different exponents β and then approximate the evidence for model $\mathcal{H}$ as

$$\ln p(d|\mathcal{H}) = \int_0^1 \langle \ln p(d|\theta) \rangle_\beta d\beta \approx \frac{1}{N_\beta} \sum_\beta \langle \ln p(d|\theta) \rangle_\beta, \quad (\text{B1})$$

where all likelihoods and posteriors are computed within model $\mathcal{H}$, θ denotes all of the model’s parameters, and the expectation $\langle \ln p(d|\theta) \rangle_\beta$ is approximated by MCMC with respect to the posterior $p_\beta(\theta|d) \propto p(d|\theta, \mathcal{H})^\beta p(\theta, \mathcal{H})$. The inverse temperatures β are spaced geometrically, as is the default in PTMCMC.

To compare nested models, which differ by “freezing” a subset of parameters, we also use the Savage–Dickey density ratio (Dickey 1971): if models $\mathcal{H}$ and $\mathcal{H}_0$ differ by the fact that (say) θ_0 is frozen to 0 in the latter, then $p(d|\mathcal{H}_0)/p(d|\mathcal{H}) = p(\theta_0 = 0|d, \mathcal{H})/p(\theta_0 = 0|\mathcal{H})$.

When comparing *disjoint* models with different likelihoods (e.g., HD vs. CURN), we use product-space sampling (Carlin & Chib 1995; Godsill 2001). This method treats model comparison as a parameter estimation problem, where we sample the union of the unique parameters of all models, plus a model-indexing parameter that activates the relevant likelihood function and parameter space of one of the submodels. Bayes factors are then obtained by counting how often the model index falls in each activation region and taking ratios of those counts.

In some situations, it can be difficult to sample a computationally expensive model directly. In these cases, we sample a computationally cheaper approximate distribution and reweight those posterior samples to estimate the posterior for the computationally expensive model (Hourihane et al. 2023). The reweighted posterior can be used in the thermodynamic integration or Savage–Dickey methods. In addition, the mean of the weights yields the Bayes factor between the expensive and approximate models, which may be of direct interest (e.g., HD can be approximated by CURN). We estimate Bayes factor uncertainties using bootstrapping and, for product-space sampling, with the Markov model techniques of Cornish & Littenberg (2015) and Heck et al. (2019).

Table 1
Prior Distributions Used in All Analyses Performed in This Paper

Parameter	Description	Prior	Comments
White Noise			
E_k	EFAC per back-end/receiver system	Uniform [0, 10]	Single-pulsar analysis only
Q_k [s]	EQUAD per back-end/receiver system	Log-uniform [−8.5, −5]	Single-pulsar analysis only
J_k [s]	ECORR per back-end/receiver system	Log-uniform [−8.5, −5]	Single-pulsar analysis only
Intrinsic Red Noise			
A_{red}	Red-noise power-law amplitude	Log-uniform [−20, −11]	One parameter per pulsar
γ_{red}	Red-noise power-law spectral index	Uniform [0, 7]	One parameter per pulsar
All Common Processes, Free Spectrum			
ρ_i [s ²]	Power-spectrum coefficients at $f = i/T$	Log-uniform in ρ_i [−18, −8]	One parameter per frequency
All Common Processes, Power-law Spectrum			
A	Common-process strain amplitude	Log-uniform [−18, −14] ($\gamma = 13/3$)	One parameter for PTA
γ	Common-process power-law spectral index	Log-uniform [−18, −11] (γ varied)	One parameter for PTA
		Delta function ($\gamma = 13/3$)	Fixed
		Uniform [0, 7]	One parameter for PTA
All Common Processes, Broken Power-law Spectrum			
A	Broken power-law amplitude	Log-uniform [−18, −11]	One parameter for PTA
γ	Broken power-law low-frequency spectral index	Uniform [0, 7]	One parameter per PTA
δ	Broken power-law high-frequency spectral index	Delta function ($\delta = 0$)	Fixed
f_{bend} [Hz]	Broken power-law bend frequency	Log-uniform [−8.7, −7]	One parameter for PTA
ℓ	Broken power-law high-frequency transition sharpness	Delta function ($\ell = 0.1$)	Fixed
All Common Processes, t-process Spectrum			
A	Power-law amplitude	Log-uniform [−18, −11]	One parameter for PTA
γ	Power-law spectral index	Uniform [0, 7]	One parameter per PTA
x_i	Modification factor	Inverse gamma distribution	One parameter per frequency
All Common Processes, Turnover Spectrum			
A	Turnover power-law amplitude	Log-uniform [−18, −11]	One parameter for PTA
γ	Turnover power-law high-frequency spectral index	Uniform [0, 7]	One parameter per PTA
κ	Turnover power-law low-frequency spectral index	Uniform [0, 7]	One parameter per PTA
f_0 [Hz]	Turnover power-law bend frequency	Log-uniform [−9, −7]	One parameter for PTA
All Common Processes, Cross-correlation Spline Model			
y	Normalized cross-correlation values at spline knots (10^{-3} , 25, 49.3, 82.5, 121.8, 150, 180) ^o	Uniform [−0.9, 0.9]	Seven parameters for PTA

Appendix C Broken Power-law Model

As shown in NG12gwb, the simultaneous Bayesian estimation of white measurement noise and of red-noise processes described by power laws biases the recovery of the spectral index of the latter (Lam et al. 2017; Hazboun et al. 2019). Just as in NG12gwb and Antoniadis et al. (2022), we impose a high-frequency cutoff on the red-noise processes. To choose the cutoff frequency, we perform inference on our data with a CURN^γ model modified so that the common process has power spectral density

$$S(f) = \frac{A^2}{12\pi^2} \left(\frac{f}{f_{\text{ref}}} \right)^{-\gamma} \left[1 + \left(\frac{f}{f_{\text{break}}} \right)^{1/\ell} \right]^{\ell\gamma} f_{\text{ref}}^{-3}, \quad (\text{C1})$$

and then set the cutoff to the MAP f_{break} . Equation (C1) is fairly generic, allowing for separate spectral indices at low (γ) and high (δ) frequencies. The break frequency f_{break} dictates where the broken power law changes spectral index, while ℓ (which we set to 0.1) controls the smoothness of the transition.

The marginal posterior for f_{break} , obtained in the factorized-likelihood approximation using the techniques of

Lamb et al. (2023), has a median and 90% credible region of $3.2^{+5.4}_{-1.2} \times 10^{-8}$ Hz and a MAP value of 2.75×10^{-8} Hz. The latter is close to $f_{14} = 14/T$ in our frequency basis (with T the total span of the data set), so we use 14 frequencies to model common-spectrum noise processes (see Section 2 and NG12gwb).

Appendix D t -process Spectrum Model

The free-spectrum analysis of our data (Section 5.2 and Figure 6) shows that the frequency bins at f_1 , f_6 , f_7 , and f_8 appear to be in tension with a pure power law, skewing the estimation of γ and reducing the HD^{13/3} versus CURN^{13/3} Bayes factor. Assuming that those frequency components reflect unmodeled systematics or stronger-than-expected statistical fluctuations, we can make our inference more robust to such outliers with a “fuzzy” power-law model that allows the individual Φ_i to vary more freely around their expected values. To wit, we introduce the t -process spectrum (TPS)

$$\Phi_{\text{TPS},i} = x_i \Phi_{\text{powerlaw},i} \quad \text{with} \quad x \sim \text{invgamma}(x_i; 1, 1), \quad (\text{D1})$$

where $\Phi_{\text{powerlaw},i}$ follows Equation (6) and x follows the inverse gamma distribution with parameters $\alpha = \beta = 1$; the resulting

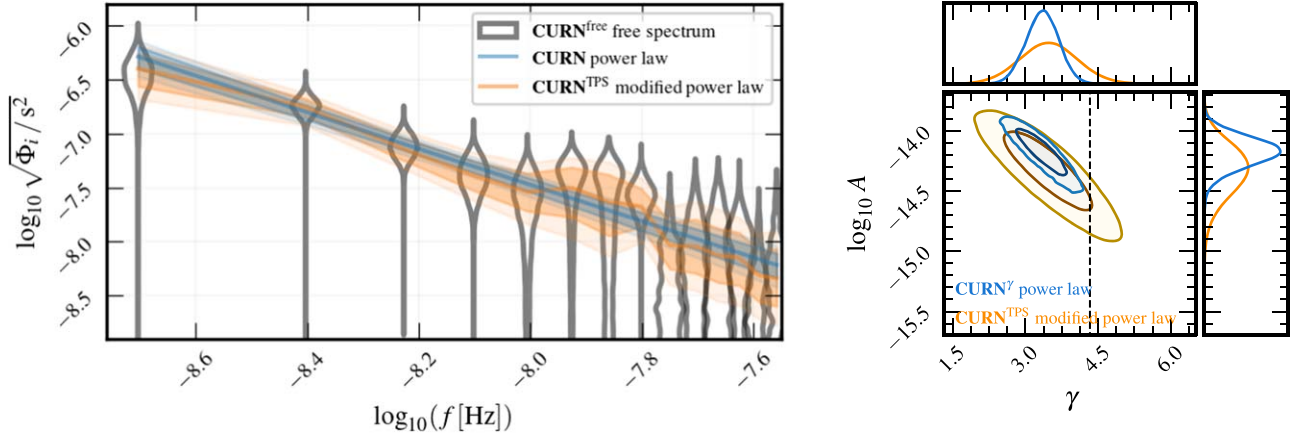


Figure 13. Power-law (CURN^{power law}; blue) and t -process power-law (CURN^{TPS}; orange) spectral posteriors. Left: reconstructed spectra, compared to free-spectral bin-variance posteriors (CURN^{TPS}; violin plots). Right: joint ($\log_{10} A$, γ) posteriors. The “fuzzy” t -process allows local deviations from a perfect power law, producing wider constraints that are more consistent with $\gamma = 13/3$ (dashed line).

Gaussian mixture yields a Student’s t -distribution for the $\Phi_{\text{TPS}, i}^{\text{TPS}}$. Figure 13 shows CURN^{power law} power-law posteriors and CURN^{TPS} modified power-law posteriors, obtained in the factorized-likelihood approximation (Taylor et al. 2022; Lamb et al. 2023) and compared to CURN^{free} bin variances. The TPS model is spread more widely and deviates from the perfect power law at bins f_1 , f_6 , f_7 , and f_8 , as expected. The right panel of Figure 13 shows the joint $\log_{10} A$, γ posteriors for CURN^{power law} and CURN^{TPS}. The latter is more consistent with steeper power laws, and it includes $\gamma = 13/3$ at 1σ credibility.

Appendix E Turnover Model

The final parameterized spectral model that we investigate is motivated by the idea that the dynamics of SMBHBs are influenced by their environments at subparsec separations (Armitage & Natarajan 2002; Sesana et al. 2004; Merritt & Milosavljević 2005). These interactions affect binary evolution and the resulting spectrum of the GWB. The process of bringing two SMBHBs together after galaxy mergers involves a complex chain of interactions: despite significant theoretical work, the lack of observational constraints makes it difficult to draw any conclusions. PTAs, however, provide a unique opportunity to probe the timescale over which two SMBHBs evolve from the merger of their galaxies to a bound binary that produces GW signals in the PTA sensitivity band.

When dynamical interactions dominate orbital evolution, binaries will traverse the GW spectrum more quickly, reducing GW emission compared to a GW-driven inspiral. This kind of behavior is captured by the *turnover model* (Sampson et al. 2015):

$$S(f) = \frac{A^2}{12\pi^2} \left(\frac{f}{f_{\text{ref}}} \right)^{-\gamma} \left[1 + \left(\frac{f_0}{f} \right)^\kappa \right]^{-1} f_{\text{ref}}^{-3}. \quad (\text{E1})$$

This is qualitatively similar to the broken power law discussed earlier, except that here f_0 represents the GW frequency at which typical binary evolution transitions from environmentally dominated (at lower frequencies and wider separations) to GW dominated (at higher frequencies and smaller separations). The parameter κ controls the shape of the spectrum below f_0 and depends on the orbital evolution mechanism. Note that the

actual turning point of the spectrum is not at f_0 but at $f_{\text{bend}} = f_0 \times (3\kappa/4 - 1)^{1/\kappa}$ (NG9gwb).

Applying this model to our data, we find hints of departures from a pure power law: the transition frequency f_0 lies below 10 nHz with 65% credibility, while the bend frequency lies below 10 nHz with 75% credibility. Nevertheless, Bayesian comparison of this CURN^{turnover} model with CURN^{power law} reports an inconclusive Bayes factor of 1.46 ± 0.02 in favor of CURN^{turnover}. Furthermore, the estimation of CURN^{turnover} parameters is sensitive to DM modeling (see Section 5.1). While the spectra are broadly consistent whether we use DMX or DMGP to model DM fluctuations, there are differences in the power at certain frequencies that lead to differences in the turnover parameters. This is discussed in greater detail in Agazie et al. (2023b).

Appendix F Sky Scrambles

In the sky scramble method (Cornish & Sampson 2016), interpulsar correlations are analyzed as if the pulsars occupied random sky positions, with the purpose of creating a background distribution of PTA detection statistics for null hypothesis testing, as an alternative to phase shifts (Taylor et al. 2017; see Sections 3 and 4). If a correlated signal is present in the data, phase shifts and sky scrambles actually test *different* null hypotheses: phase shifts test the hypothesis that no interpulsar correlations are present, while sky scrambles assume that interpulsar correlations are present at the level measured in the data but test the hypothesis that these correlations have no dependence on angular separation.

As is the convention in the literature, we require that scrambled overlap reduction functions (ORFs) be independent of each other and of the unscrambled ORF using a match statistic,

$$\bar{M} = \frac{\sum_{a,b \neq a} \Gamma_{ab} \Gamma'_{ab}}{\sqrt{(\sum_{a,b \neq a} \Gamma_{ab} \Gamma_{ab}) (\sum_{a,b \neq a} \Gamma'_{ab} \Gamma'_{ab})}}, \quad (\text{F1})$$

where Γ_{ab} and Γ'_{ab} are two different ORFs. For the sky scrambles used in our analysis, the scrambled ORFs have $\bar{M} < 0.1$ with respect to the unscrambled ORF and $\bar{M} < 0.17$ with respect to each other. We generate 10,000 sky scrambles,

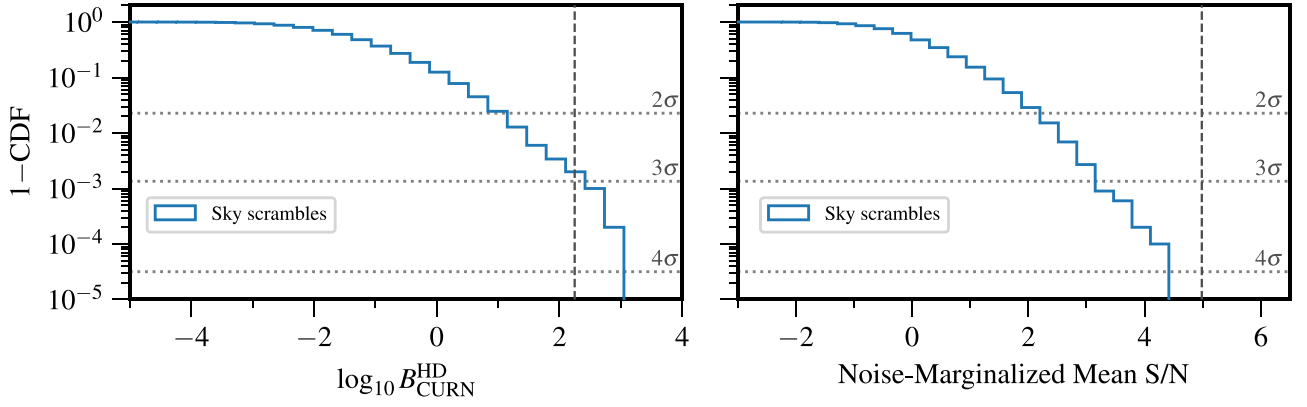


Figure 14. Empirical background distribution of $\text{HD}^\gamma\text{-to-CURN}^\gamma$ Bayes factor (left; see Section 3) and noise-marginalized optimal statistic (right; see Section 4), as computed in 5000 sky scrambles, which erases the dependence of interpulsar correlations on the angular separation between the pulsars. Dotted lines indicate Gaussian-equivalent 2σ , 3σ , and 4σ thresholds. The dashed vertical lines indicate the values of the detection statistics for the unscrambled data set. We find $p = 1.6 \times 10^{-3}$ ($\sim 3\sigma$) for the Bayesian analysis and $p < 10^{-4}$ ($> 3\sigma$) for the optimal statistic analysis.

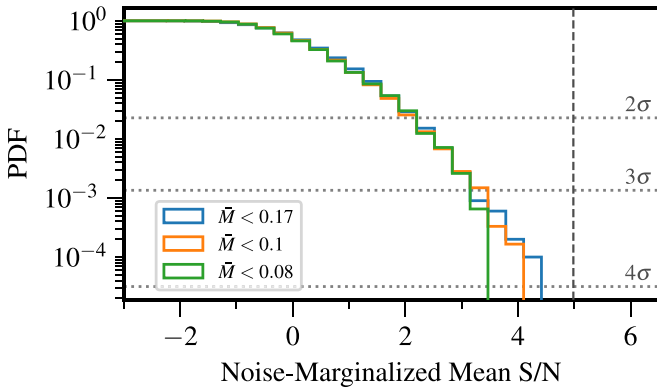


Figure 15. Comparison between empirical background distributions for the noise-marginalized optimal statistic, as computed by the sky scramble technique. We show distributions computed using a match threshold of $\bar{M} < 0.17$ (blue), $\bar{M} < 0.1$ (orange), and $\bar{M} < 0.08$ (green). Dotted lines indicate Gaussian-equivalent 2σ , 3σ , and 4σ thresholds. The dashed vertical lines indicate the values of the detection statistics for the unscrambled data set. We find little difference between the background distributions computed using different match thresholds, modulo the fact that imposing a smaller threshold yields fewer sky scrambles, which limits the precision to which the p -value can be measured.

owing to the difficulty in obtaining large numbers of scrambled ORFs that satisfy the match threshold; because of limitations of computational resources, we obtain our detection statistics for 5000 of those ORFs. Figure 14 shows the resulting background distributions for the $\text{HD}^\gamma\text{-to-CURN}^\gamma$ Bayes factor (left panel) and the optimal statistic S/N (right panel). The Bayes factors exceed the observed value in 8 of the 5000 sky scrambles ($p = 1.6 \times 10^{-3}$), while none of the sky scrambles have noise-marginalized mean S/N greater than observed ($p < 10^{-4}$).

We note that the null distribution recovered by the sky scrambles is not very sensitive to the choice of match threshold for $|\bar{M}| \lesssim 0.2$. Figure 15 compares the null distributions when the match threshold for all ORFs with each other and with the unscrambled ORF is set to $|\bar{M}| < 0.17$ (blue), $|\bar{M}| < 0.1$ (orange), and $|\bar{M}| < 0.08$ (green). There is very little difference among the distributions; however, imposing a smaller threshold means that fewer sky scrambles can be used (6043 with $|\bar{M}| < 0.1$ and 1534 with $|\bar{M}| < 0.08$, compared to 10,000 with $|\bar{M}| < 0.17$), which limits the precision with which the p -value can be measured. We find no evidence that the recovered null

distribution is biased when including sky scrambles with matches up to 0.17.

Appendix G Multiple-correlation Optimal Statistic

The multiple-correlation optimal statistic (MCOS; Sardesai & Vigeland 2023) fits the interpulsar correlation coefficients ρ_{ab} with a linear model that includes multiple components with different correlation patterns, but with the same spectral shape. The linear model coefficients are the squared amplitudes of the components. Within such a model, the significance of each component can be quoted as an S/N given by its best-fit coefficient divided by the fit error. Just as for the noise-marginalized optimal statistic (Vigeland et al. 2018), the posterior distribution of pulsar noise parameters induces a distribution of MCOS statistics.

We fit the 15 yr data with models that include HD, monopole, and dipole-correlated components in various combinations. Table 2 lists the noise-marginalized amplitude estimates and S/N for all models. The goodness of fit of the models can be compared using the Akaike Information Criterion (AIC; Akaike 1998):

$$\text{AIC} = 2k + \chi^2, \quad (\text{G1})$$

where k is the number of model parameters and χ^2 is the fit's chi-squared, computed without accounting for GW-induced ρ_{ab} correlations. (This can be thought of as a Bayes factor analog, with the factor of $2k$ imposing an Occam penalty.) The relative probability of a model compared to the most-favored model is then given by

$$p(\text{AIC}) = \exp[(\text{AIC}_{\min} - \text{AIC})/2], \quad (\text{G2})$$

where $\text{AIC}_{\min}$ is the minimum AIC across all models.

Table 2 lists the AIC probabilities, computed by averaging the AIC of each model over pulsar noise parameters. The HD-correlated model is preferred among the models with a single correlated process. The models with both HD and monopole correlations are preferred among all models: for a model with HD and monopole correlations, we find S/Ns of 3.4 ± 0.8 for HD correlations and 2.9 ± 0.8 for monopolar correlations (see Figure 16), while for a model with HD, monopole, and dipole correlations, we find S/Ns of 2.9 ± 0.6 for HD correlations, 2.4 ± 0.6 for monopole correlations, and 0.6 ± 0.4 for dipole

Table 2
Multiple-correlation Optimal Statistic Best-fit Coefficients $\hat{A}_k^2$, S/Ns, and AIC Probabilities

Model	HD Correlations		Monopole Correlations		Dipole Correlations		$p(\text{AIC})$
	$\hat{A}_{\text{HD}}^2$	S/N	$\hat{A}_{\text{mo}}^2$	S/N	Mean $\hat{A}_{\text{di}}^2$	S/N	
HD only	$6.8(9) \times 10^{-30}$	4(1)	...	...	...	...	3×10^{-2}
Monopole only	...	...	$1.1(1) \times 10^{-30}$	4(1)	...	...	6×10^{-3}
Dipole only	...	...	...	...	$1.5(3) \times 10^{-30}$	4(1)	8×10^{-4}
HD + monopole	$5.5(8) \times 10^{-30}$	3.4(8)	$8(1) \times 10^{-31}$	2.9(8)	...	...	1
HD + dipole	$5.5(8) \times 10^{-30}$	3.2(7)	...	...	$8(2) \times 10^{-31}$	1.7(7)	6×10^{-2}
Monopole + dipole	...	...	$8(1) \times 10^{-31}$	2.7(7)	$9(2) \times 10^{-31}$	1.9(6)	1×10^{-2}
HD + monopole + dipole	$5.1(8) \times 10^{-30}$	2.9(6)	$7(1) \times 10^{-31}$	2.4(6)	$3(2) \times 10^{-31}$	0.6(4)	0.48

Note. All values were computed for the 15 yr data set, assuming a power-law power spectral density using the 14 lowest-frequency components. Here $\hat{A}_k^2$, S/N, and AIC are marginalized over pulsar noise parameters with fixed $\gamma = 13/3$. The numbers in parentheses represent the mean least-squares errors for the $\hat{A}_k^2$ coefficients and standard deviations over noise parameter posteriors for S/Ns. We compute $p(\text{AIC})$ with respect to the model with the lowest mean AIC (i.e., HD + monopole).

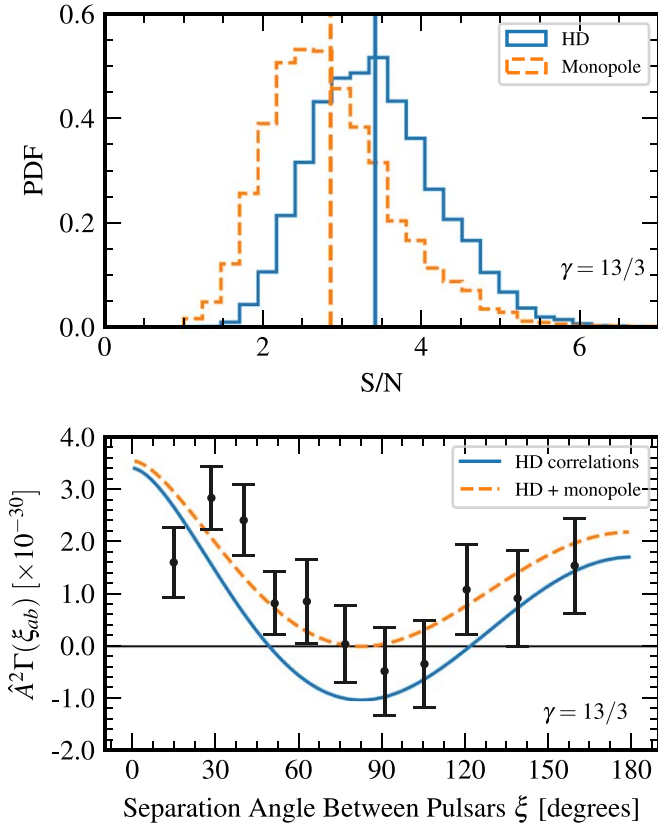


Figure 16. Results of the MCOS analysis, which prefers a model including both HD and monopole correlations. Top: MCOS S/N for HD correlations (solid blue) and monopole correlations (dashed orange), marginalized over CURN^{13/3} noise parameter posteriors. The vertical lines indicate the mean S/Ns. We find S/Ns of 3.4 ± 0.8 for HD correlations and 2.9 ± 0.8 for monopole correlations. Bottom: binned cross-correlations ρ_{ab} (black error bars), computed with MAP noise parameters from a CURN^{13/3} run. The solid blue and dashed orange curves show best-fit HD and HD+monopole correlation patterns, corresponding to $\hat{A}^2 = 6.8 \times 10^{-30}$ and to $\hat{A}_{\text{HD}}^2 = 5.5 \times 10^{-30}$, $\hat{A}_{\text{monopole}}^2 = 8 \times 10^{-31}$, respectively. The monopolar component accounts for the vertical shift of the cross-correlations with respect to the HD curve. We use the standard version of the optimal statistic that does not include interpulsar correlations to compute ρ_{ab} , so the points and errors do not match those shown in Figure 1(c).

correlations (means $\pm$ standard deviations across noise parameter posteriors). The statistical significance of these S/Ns can be quantified empirically using simulations of 15 yr–like data

sets (see Appendix H.1), which report p -values $< 10^{-2}$ and $\approx 4 \times 10^{-2}$ for the observed mean HD and monopole statistics across data replications with no spatially correlated injections.

As discussed in Sardesai & Vigeland (2023), the optimal statistic and the MCOS are metrics of the *apparent* spatial correlation pattern of the data, but they have a limited ability to identify its actual source. That is because a real HD signal may also excite the monopole optimal statistic and the monopole component of the MCOS; conversely, a real monopolar signal may also excite the HD optimal statistic and the HD component of the MCOS; and so on. The S/Ns quoted in Table 2 quantify how often we would expect to measure the observed value of the optimal statistic if only uncorrelated noise is present, but they do not describe how often one type of correlated noise would produce a given value of the optimal statistic for a different type of correlation. This effect can be characterized using simulations (see Appendix H.1), which report a p -value of 0.11 for the observed mean monopole statistic when an HD-correlated signal with the MAP 15 yr amplitude is included in the simulated data sets. We conclude that there are some indications of a possible monopole-correlated signal in the data with S/N comparable to but smaller than the S/N for HD correlations; however, from simulations we conclude that it is possible for such a signal to appear in an MCOS analysis if only an HD-correlated stochastic process is present.

Appendix H Multiple-correlation Optimal Statistic Simulations

In this appendix we obtain the distribution of the MCOS over an ensemble of simulated data sets, with the goal of characterizing the probability that the observed S/Ns could have been produced by pulsar noise alone, or by a GWB with HD correlations. Unlike our Bayesian analysis, the MCOS prefers a model that includes both HD and monopolar components. Hence, we are especially interested in asking how frequently we may expect the observed MCOS monopole if the data contain only the GWB. In Appendices H.1 and H.2 we present two different types of simulations: “astrophysical,” where we generate synthetic data with MAP noise parameters inferred from the 15 yr data set, both with and without the GWB; and “model checking,” where we create data replications following the HD^{13/3} posteriors for the real data set. Note that neither simulation attempts to account for the monochromatic character of the putative monopolar signal (see Section 5.2).

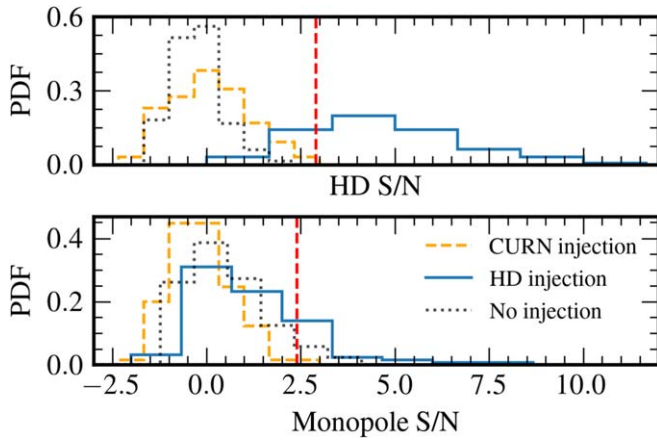


Figure 17. Top: MCOS HD S/N values recovered in the three simulations described in Appendix H.1, compared to the MCOS HD S/N measured in the real data set (vertical dashed red line), which has p -values of $<10^{-2}$ for simulations (i) and (iii) and 0.64 for simulation (ii). Bottom: MCOS monopole S/N values recovered in the three simulations, compared to the real-data MCOS monopole S/N (vertical dashed red line), which has p -values of 4×10^{-2} , 1.1×10^{-1} , and $<10^{-2}$ for simulations (i), (ii), and (iii), respectively.

H.1. Astrophysical Simulations

Following Pol et al. (2021), we generate simulated data sets adopting MAP pulsar noise parameters obtained from the real data independently for each pulsar; these “noise runs” include an additional power-law process to reduce contamination between the putative GWB and the pulsars’ intrinsic red noise (Taylor et al. 2022). We produce 100 realizations each of three different simulations: (i) injecting no spatially correlated power-law GWB or excess uncorrelated common-spectrum noise; (ii) injecting a spatially correlated power-law GWB with amplitude 2.7×10^{-15} and spectral index 13/3; and (iii) injecting no GWB or common-spectrum noise, but omitting the additional power-law process in the estimation of intrinsic pulsar noise, with the goal of testing how often excess common-spectrum noise is recognized as a spatially correlated GWB.

We compute HD + monopole + dipole MCOS S/Ns for all synthetic data sets (see Figure 17). The mean HD S/Ns observed in the real data (see Appendix G) correspond to p -values of $<10^{-2}$ for simulations (i) and (iii) and 0.64 for simulation (ii). The mean monopole S/Ns observed in the real data set correspond to p -values of 4×10^{-2} , 1.1×10^{-1} , and $<10^{-2}$ for simulations (i), (ii), and (iii), respectively. We conclude that it is unlikely that we would measure HD correlations at the level observed in real data when no correlated signal is present (simulation (i)) or when only uncorrelated common-spectrum red noise is present (simulation (iii)). In addition, the HD S/Ns obtained from an HD-correlated GWB injection (simulation (ii)) are fully consistent with the S/N observed in real data. By contrast, the observed monopole S/N could have been produced by intrinsic pulsar noise alone, or by a real HD signal.

H.2. Model-checking Simulations

In Appendix H.1 we have tackled the question of monopole S/N significance using simulations based on real-data MAP estimates η^{MAP} of pulsar noise and GW parameters. In this appendix we adopt a procedure with a stronger Bayesian flavor,

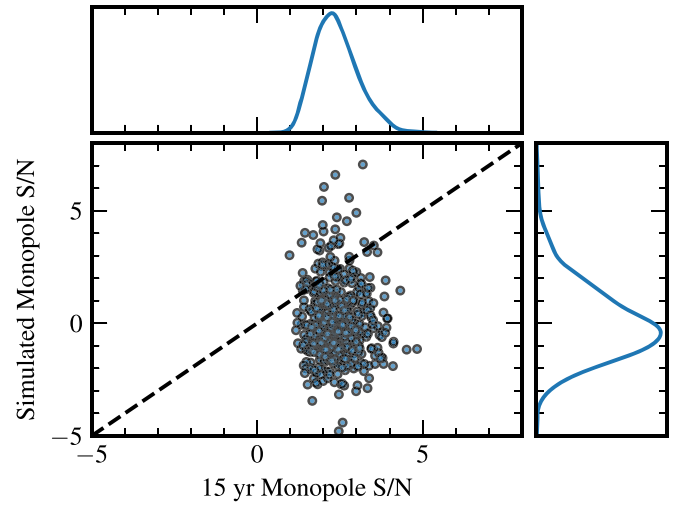


Figure 18. Distribution of real-data and replicated MCOS monopole S/Ns. Each point represents a draw $\eta^{(k)}$ from $\text{HD}^{13/3}$ posterior, which is used to simulate $\delta t^{\text{sim},(k)}$ and to compute both S/Ns. The replicated monopole S/N is greater for 11% of the simulations.

evaluating the MCOS on a population of *data replications* created using $\text{HD}^{13/3}$ as a generative model with noise hyperparameters η drawn from the $\text{HD}^{13/3}$ real-data posterior. This can be seen also as a Bayesian model-checking exercise (Gelman et al. 1996, 2013): if we find that the summary statistic of interest (the monopole MCOS) has a much more extreme value in real data than in data replications, we should suspect that the data model (here $\text{HD}^{13/3}$) is missing something.

We perform the test by drawing 500 parameter vectors $\{\eta^{(k)}\}$ from the $\text{HD}^{13/3}$ real-data posterior; for each $\eta^{(k)}$ we simulate a data set $\delta t^{\text{sim},(k)} \sim p(\delta t | \eta^{(k)})$ and compare $\text{MCOS}(\delta t^{\text{sim},(k)}; \eta^{(k)})$ with $\text{MCOS}(\delta t; \eta^{(k)})$. Our notation emphasizes the dependence of the MCOS on the pulsar noise parameters through the $\mathbf{P}$ matrices in Equation (9). Figure 18 shows the resulting distribution of monopole S/Ns. The replicated monopole S/N is greater than its observed counterpart for 11% of the draws. Thus, it is plausible that the MCOS could measure the observed monopole S/N in data that contain only an HD-correlated GWB. Conversely, the observed monopole S/N does not by itself suggest that $\text{HD}^{13/3}$ is misspecified.

ORCID iDs

Gabriella Agazie <https://orcid.org/0000-0001-5134-3925>
 Akash Anumalapudi <https://orcid.org/0000-0002-8935-9882>
 Anne M. Archibald <https://orcid.org/0000-0003-0638-3340>
 Paul T. Baker <https://orcid.org/0000-0003-2745-753X>
 Bence Bécsy <https://orcid.org/0000-0003-0909-5563>
 Laura Blecha <https://orcid.org/0000-0002-2183-1087>
 Adam Brazier <https://orcid.org/0000-0001-6341-7178>
 Paul R. Brook <https://orcid.org/0000-0003-3053-6538>
 Sarah Burke-Spolaor <https://orcid.org/0000-0003-4052-7838>
 Maria Charisi <https://orcid.org/0000-0003-3579-2522>
 Shami Chatterjee <https://orcid.org/0000-0002-2878-1502>
 Katerina Chatziioannou <https://orcid.org/0000-0002-5833-413X>
 Siyuan Chen <https://orcid.org/0000-0002-3118-5963>
 Tyler Cohen <https://orcid.org/0000-0001-7587-5483>

- James M. Cordes  <https://orcid.org/0000-0002-4049-1882>
 Neil J. Cornish  <https://orcid.org/0000-0002-7435-0869>
 Fronefield Crawford  <https://orcid.org/0000-0002-2578-0360>
 H. Thankful Cromartie  <https://orcid.org/0000-0002-6039-692X>
 Kathryn Crowter  <https://orcid.org/0000-0002-1529-5169>
 Curt J. Cutler  <https://orcid.org/0000-0002-2080-1468>
 Megan E. DeCesar  <https://orcid.org/0000-0002-2185-1790>
 Paul B. Demorest  <https://orcid.org/0000-0002-6664-965X>
 Timothy Dolch  <https://orcid.org/0000-0001-8885-6388>
 Elizabeth C. Ferrara  <https://orcid.org/0000-0001-7828-7708>
 William Fiore  <https://orcid.org/0000-0001-5645-5336>
 Emmanuel Fonseca  <https://orcid.org/0000-0001-8384-5049>
 Gabriel E. Freedman  <https://orcid.org/0000-0001-7624-4616>
 Nate Garver-Daniels  <https://orcid.org/0000-0001-6166-9646>
 Peter A. Gentile  <https://orcid.org/0000-0001-8158-683X>
 Joseph Glaser  <https://orcid.org/0000-0003-4090-9780>
 Deborah C. Good  <https://orcid.org/0000-0003-1884-348X>
 Kayhan Gültekin  <https://orcid.org/0000-0002-1146-0198>
 Jeffrey S. Hazboun  <https://orcid.org/0000-0003-2742-3321>
 Sophie Hourihane  <https://orcid.org/0000-0002-9152-0719>
 Ross J. Jennings  <https://orcid.org/0000-0003-1082-2342>
 Aaron D. Johnson  <https://orcid.org/0000-0002-7445-8423>
 Megan L. Jones  <https://orcid.org/0000-0001-6607-3710>
 Andrew R. Kaiser  <https://orcid.org/0000-0002-3654-980X>
 David L. Kaplan  <https://orcid.org/0000-0001-6295-2881>
 Luke Zoltan Kelley  <https://orcid.org/0000-0002-6625-6450>
 Matthew Kerr  <https://orcid.org/0000-0002-0893-4073>
 Joey S. Key  <https://orcid.org/0000-0003-0123-7600>
 Nima Laal  <https://orcid.org/0000-0002-9197-7604>
 Michael T. Lam  <https://orcid.org/0000-0003-0721-651X>
 William G. Lamb  <https://orcid.org/0000-0003-1096-4156>
 Natalia Lewandowska  <https://orcid.org/0000-0003-0771-6581>
 Tyson B. Littenberg  <https://orcid.org/0000-0002-9574-578X>
 Tingting Liu  <https://orcid.org/0000-0001-5766-4287>
 Andrea Lommen  <https://orcid.org/0000-0003-4137-7536>
 Duncan R. Lorimer  <https://orcid.org/0000-0003-1301-966X>
 Jing Luo  <https://orcid.org/0000-0001-5373-5914>
 Ryan S. Lynch  <https://orcid.org/0000-0001-5229-7430>
 Chung-Pei Ma  <https://orcid.org/0000-0002-4430-102X>
 Dustin R. Madison  <https://orcid.org/0000-0003-2285-0404>
 Alexander McEwen  <https://orcid.org/0000-0001-5481-7559>
 James W. McKee  <https://orcid.org/0000-0002-2885-8485>
 Maura A. McLaughlin  <https://orcid.org/0000-0001-7697-7422>
 Natasha McMann  <https://orcid.org/0000-0002-4642-1260>
 Bradley W. Meyers  <https://orcid.org/0000-0001-8845-1225>
 Patrick M. Meyers  <https://orcid.org/0000-0002-2689-0190>
 Chiara M. F. Mingarelli  <https://orcid.org/0000-0002-4307-1322>
 Andrea Mitridate  <https://orcid.org/0000-0003-2898-5844>
 Priyamvada Natarajan  <https://orcid.org/0000-0002-5554-8896>
 Cherry Ng  <https://orcid.org/0000-0002-3616-5160>
 David J. Nice  <https://orcid.org/0000-0002-6709-2566>
 Stella Koch Ocker  <https://orcid.org/0000-0002-4941-5333>
 Ken D. Olum  <https://orcid.org/0000-0002-2027-3714>
 Timothy T. Pennucci  <https://orcid.org/0000-0001-5465-2889>
 Benetge B. P. Perera  <https://orcid.org/0000-0002-8509-5947>
 Polina Petrov  <https://orcid.org/0000-0001-5681-4319>
 Nihan S. Pol  <https://orcid.org/0000-0002-8826-1285>
 Henri A. Radovan  <https://orcid.org/0000-0002-2074-4360>
 Scott M. Ransom  <https://orcid.org/0000-0001-5799-9714>
 Paul S. Ray  <https://orcid.org/0000-0002-5297-5278>
 Joseph D. Romano  <https://orcid.org/0000-0003-4915-3246>
 Shashwat C. Sardesai  <https://orcid.org/0009-0006-5476-3603>
 Ann Schmiedekamp  <https://orcid.org/0000-0003-4391-936X>
 Carl Schmiedekamp  <https://orcid.org/0000-0002-1283-2184>
 Kai Schmitz  <https://orcid.org/0000-0003-2807-6472>
 Levi Schult  <https://orcid.org/0000-0001-6425-7807>
 Brent J. Shapiro-Albert  <https://orcid.org/0000-0002-7283-1124>
 Xavier Siemens  <https://orcid.org/0000-0002-7778-2990>
 Joseph Simon  <https://orcid.org/0000-0003-1407-6607>
 Magdalena S. Siwek  <https://orcid.org/0000-0002-1530-9778>
 Ingrid H. Stairs  <https://orcid.org/0000-0001-9784-8670>
 Daniel R. Stinebring  <https://orcid.org/0000-0002-1797-3277>
 Kevin Stovall  <https://orcid.org/0000-0002-7261-594X>
 Jerry P. Sun  <https://orcid.org/0000-0002-7933-493X>
 Abhimanyu Susobhanan  <https://orcid.org/0000-0002-2820-0931>
 Joseph K. Swiggum  <https://orcid.org/0000-0002-1075-3837>
 Stephen R. Taylor  <https://orcid.org/0000-0001-8217-1599>
 Jacob E. Turner  <https://orcid.org/0000-0002-2451-7288>
 Caner Unal  <https://orcid.org/0000-0001-8800-0192>
 Michele Vallisneri  <https://orcid.org/0000-0002-4162-0033>
 Rutger van Haasteren  <https://orcid.org/0000-0002-6428-2620>
 Sarah J. Vigeland  <https://orcid.org/0000-0003-4700-9072>
 Haley M. Wahl  <https://orcid.org/0000-0001-9678-0299>
 Caitlin A. Witt  <https://orcid.org/0000-0002-6020-9274>
 Olivia Young  <https://orcid.org/0000-0002-0883-0688>

References

- Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, *PhRvL*, **116**, 061102
 Afzal, A., Agazie, G., Anumalapudi, A., et al. 2023, *ApJL*, **951**, L11
 Agazie, G., Alam, M. F., Anumalapudi, A., et al. 2023a, *ApJL*, **951**, L9
 Agazie, G., Anumalapudi, A., Archibald, A. M., et al. 2023b, *ApJL*, submitted
 Agazie, G., Anumalapudi, A., Archibald, A. M., et al. 2023c, *ApJL*, **951**, L10
 Aggarwal, K., Arzoumanian, Z., Baker, P. T., et al. 2019, *ApJ*, **880**, 116
 Akaike, H. 1998, in *Information Theory and an Extension of the Maximum Likelihood Principle*, ed. E. Parzen, K. Tanabe, & G. Kitagawa (New York: Springer), 199
 Alam, M. F., Arzoumanian, Z., Baker, P. T., et al. 2021a, *ApJS*, **252**, 4
 Alam, M. F., Arzoumanian, Z., Baker, P. T., et al. 2021b, *ApJS*, **252**, 5
 Allen, B. 2023, *PhRvD*, **107**, 043018
 Allen, B., & Romano, J. D. 2022, arXiv:2208.07230
 Anholm, M., Ballmer, S., Creighton, J. D. E., Price, L. R., & Siemens, X. 2009, *PhRvD*, **79**, 084030
 Antoniadis, J., Arzoumanian, Z., Babak, S., et al. 2022, *MNRAS*, **510**, 4873
 Armitage, P. J., & Natarajan, P. 2002, *ApJL*, **567**, L9
 Arzoumanian, Z., Baker, P. T., Blecha, L., et al. 2023, arXiv:2301.03608
 Arzoumanian, Z., Baker, P. T., Blumer, H., et al. 2020, *ApJL*, **905**, L34

- Arzoumanian, Z., Baker, P. T., Blumer, H., et al. 2021, *ApJL*, **923**, L22
- Arzoumanian, Z., Baker, P. T., Brazier, A., et al. 2018, *ApJ*, **859**, 47
- Arzoumanian, Z., Brazier, A., Burke-Spolaor, S., et al. 2015, *ApJ*, **813**, 65
- Arzoumanian, Z., Brazier, A., Burke-Spolaor, S., et al. 2016, *ApJ*, **821**, 13
- Astropy Collaboration, Price-Whelan, A. M., Lim, A. M., et al. 2022, *ApJ*, **935**, 167
- Backer, D. C., Kulkarni, S. R., Heiles, C., Davis, M. M., & Goss, W. M. 1982, *Natur*, **300**, 615
- Bartolo, N., Bertacca, D., Caldwell, R., et al. 2022, *JCAP*, **2022**, 009
- Bécsy, B., Cornish, N. J., & Kelley, L. Z. 2022, *ApJ*, **941**, 119
- Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, *Natur*, **287**, 307
- Blumenthal, G. R., Faber, S. M., Primack, J. R., & Rees, M. J. 1984, *Natur*, **311**, 517
- Burke-Spolaor, S., Taylor, S. R., Charisi, M., et al. 2019, *A&ARv*, **27**, 5
- Caprini, C., & Figuerola, D. G. 2018, *CQGr*, **35**, 163001
- Carlin, B. P., & Chib, S. 1995, *J. R. Stat. Soc. Series B Stat. Methodol.*, **57**, 473
- Chalumeau, A., Babak, S., Petiteau, A., et al. 2022, *MNRAS*, **509**, 5538
- Chamberlin, S. J., Creighton, J. D. E., Siemens, X., et al. 2015, *PhRvD*, **91**, 044048
- Chen, S., Caballero, R. N., Guo, Y. J., et al. 2021, *MNRAS*, **508**, 4970
- Cordes, J. M. 2013, *CQGr*, **30**, 224002
- Cornish, N. J., & Littenberg, T. B. 2015, *CQGr*, **32**, 135012
- Cornish, N. J., & Sampson, L. 2016, *PhRvD*, **93**, 104047
- Cornish, N. J., & Sesana, A. 2013, *CQGr*, **30**, 224005
- Demorest, P. B. 2007, PhD thesis, Univ. California, Berkeley
- Demorest, P. B., Ferdman, R. D., Gonzalez, M. E., et al. 2013, *ApJ*, **762**, 94
- Desvignes, G., Caballero, R. N., Lentati, L., et al. 2016, *MNRAS*, **458**, 3341
- Detweiler, S. 1979, *ApJ*, **234**, 1100
- Dickey, J. M. 1971, *Ann. Math. Stat.*, **42**, 204
- Domènech, G. 2021, *Univ*, **7**, 398
- DuPlain, R., Ransom, S., Demorest, P., et al. 2008, *Proc. SPIE*, **7019**, 70191D
- Einstein, A. 1916, Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.), **1916**, 1
- Ellis, J., & van Haasteren, R. 2017, jellis18/PTMCMCSampler: Official Release v1.0.0 Zenodo, doi:10.5281/zenodo.1037579
- Ellis, J. A., Vallisneri, M., Taylor, S. R., & Baker, P. T. 2019, ENTERPRISE: Enhanced Numerical Toolbox Enabling a Robust Pulsar Inference Suite, Astrophysics Source Code Library, ascl:1912.015
- Enoki, M., & Nagashima, M. 2007, *PTHPh*, **117**, 241
- Estabrook, F. B., & Wahlquist, H. D. 1975, *GReGr*, **6**, 439
- Ford, J. M., Demorest, P., & Ransom, S. 2010, *Proc. SPIE*, **7740**, 77400A
- Foster, R. S., & Backer, D. C. 1990, *ApJ*, **361**, 300
- Gair, J., Romano, J. D., Taylor, S., & Mingarelli, C. M. F. 2014, *PhRvD*, **90**, 082001
- Gelman, A., Carlin, J., Stern, H., et al. 2013, Bayesian Data Analysis (3rd ed.; London: Chapman and Hall)
- Gelman, A., & Meng, X.-L. 1998, *StaSc*, **13**, 163
- Gelman, A., Meng, X.-L., & Stern, H. 1996, *Stat. Sin.*, **6**, 733
- Gelman, A., & Rubin, D. B. 1992, *StaSc*, **7**, 457
- Godsill, S. J. 2001, *J. Comput. Graph. Stat.*, **10**, 230
- Goncharov, B., Reardon, D. J., Shannon, R. M., et al. 2021a, *MNRAS*, **502**, 478
- Goncharov, B., Shannon, R. M., Reardon, D. J., et al. 2021b, *ApJL*, **917**, L19
- Goncharov, B., Thrane, E., Shannon, R. M., et al. 2022, *ApJL*, **932**, L22
- Guzzetti, M. C., Bartolo, N., Liguori, M., & Matarrese, S. 2016, *NCimR*, **39**, 399
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, *Natur*, **585**, 357
- Hazboun, J., Meyers, P. M., Romano, J. D., Siemens, X., & Archibald, A. M. 2023, arXiv:2305.01116
- Hazboun, J., Romano, J., & Smith, T. 2019, *JOSS*, **4**, 1775
- Hazboun, J. S., Romano, J. D., & Smith, T. L. 2019, *PhRvD*, **100**, 104028
- Hazboun, J. S., Simon, J., Madison, D. R., et al. 2022, *ApJ*, **929**, 39
- Hazboun, J. S., Simon, J., Taylor, S. R., et al. 2020, *ApJ*, **890**, 108
- Heck, D. W., Overstall, A. M., Gronau, Q. F., & Wagenmakers, E.-J. 2019, *Stat. Comput.*, **29**, 631
- Hee, S., Handley, W. J., Hobson, M. P., & Lasenby, A. N. 2015, *MNRAS*, **455**, 2461
- Hellings, R. W., & Downs, G. S. 1983, *ApJL*, **265**, L39
- Hinton, S. R. 2016, *JOSS*, **1**, 00045
- Hobbs, G., Coles, W., Manchester, R. N., et al. 2012, *MNRAS*, **427**, 2780
- Hobbs, G., & Edwards, R. 2012, Tempo2: Pulsar Timing Package, Astrophysics Source Code Library, ascl:1210.015
- Hobbs, G., Guo, L., Caballero, R. N., et al. 2020, *MNRAS*, **491**, 5951
- Hourihaane, S., Meyers, P., Johnson, A., Chatziioannou, K., & Vallisneri, M. 2023, *PhRvD*, **107**, 084045
- Hulse, R. A., & Taylor, J. H. 1975, *ApJL*, **195**, L51
- Hunter, J. D. 2007, *CSE*, **9**, 90
- Jaffe, A. H., & Backer, D. C. 2003, *ApJ*, **583**, 616
- Jenet, F. A., Hobbs, G. B., van Straten, W., et al. 2006, *ApJ*, **653**, 1571
- Jones, M. L., McLaughlin, M. A., Lam, M. T., et al. 2017, *ApJ*, **841**, 125
- Joshi, B. C., Arumugasamy, P., Bagchi, M., et al. 2018, *JApA*, **39**, 51
- Kaiser, A. R., Pol, N. S., McLaughlin, M. A., et al. 2022, *ApJ*, **938**, 115
- Kaspi, V. M., Taylor, J. H., & Ryba, M. F. 1994, *ApJ*, **428**, 713
- Khmelnitsky, A., & Rubakov, V. 2014, *JCAP*, **2014**, 019
- Kluyver, T., Ragan-Kelley, B., Pérez, F., et al. 2016, in Positioning and Power in Academic Publishing: Players, Agents and Agendas, ed. F. Loizides & B. Schmidt (Amsterdam: IOS Press), 87
- Lam, M. T., Cordes, J. M., Chatterjee, S., et al. 2017, *ApJ*, **834**, 35
- Lam, M. T., Ellis, J. A., Grillo, G., et al. 2018, *ApJ*, **861**, 132
- Lamb, W. G., Taylor, S. R., & van Haasteren, R. 2023, arXiv:2303.15442
- Lee, K. J. 2016, in ASP Conf. Ser. 502, Frontiers in Radio Astronomy and FAST Early Sciences Symposium 2015, ed. L. Qain & D. Li (San Francisco, CA: ASP), 19
- Lentati, L., Alexander, P., Hobson, M. P., et al. 2013, *PhRvD*, **87**, 104021
- Luo, J., Ransom, S., Demorest, P., et al. 2021, *ApJ*, **911**, 45
- Magorrian, J., Tremaine, S., Richstone, D., et al. 1998, *AJ*, **115**, 2285
- Manchester, R. N., Hobbs, G., Bailes, M., et al. 2013, *PASA*, **30**, e017
- McConnell, N. J., & Ma, C.-P. 2013, *ApJ*, **764**, 184
- McLaughlin, M. A. 2013, *CQGr*, **30**, 224008
- McWilliams, S. T., Ostriker, J. P., & Pretorius, F. 2014, *ApJ*, **789**, 156
- Merritt, D., & Milosavljević, M. 2005, *LRR*, **8**, 8
- Meyers, P. M., Chatziioannou, K., Vallisneri, M., & Chua, A. J. K. 2023, arXiv:2306.05559
- Milosavljević, M., & Merritt, D. 2003, *ApJ*, **596**, 860
- Mingarelli, C. M. F., Lazio, T. J. W., Sesana, A., et al. 2017, *NatAs*, **1**, 886
- Mingarelli, C. M. F., & Mingarelli, A. B. 2018, *JPhCo*, **2**, 105002
- Mingarelli, C. M. F., & Sidery, T. 2014, *PhRvD*, **90**, 062011
- Mingarelli, C. M. F., Sidery, T., Mandel, I., & Vecchio, A. 2013, *PhRvD*, **88**, 062005
- Nay, J., Boddy, K. K., Smith, T. L., & Mingarelli, C. M. F. M. 2023, arXiv:2306.06168
- Ogata, Y. 1989, *NuMat*, **55**, 137
- Park, R. S., Folkner, W. M., Williams, J. G., & Boggs, D. H. 2021, *AJ*, **161**, 105
- Perera, B. B. P., DeCesar, M. E., Demorest, P. B., et al. 2019, *MNRAS*, **490**, 4666
- Phinney, E. S. 2001, arXiv:astro-ph/0108028
- Pirani, F. A. E. 1956, *AcPP*, **15**, 389
- Pirani, F. A. E. 2009, *GReGr*, **41**, 1215
- Pol, N., Taylor, S. R., & Romano, J. D. 2022, *ApJ*, **940**, 173
- Pol, N. S., Taylor, S. R., Kelley, L. Z., et al. 2021, *ApJL*, **911**, L34
- Rajagopal, M., & Romani, R. W. 1995, *ApJ*, **446**, 543
- Ransom, S., Brazier, A., Chatterjee, S., et al. 2019, *BAAS*, **51**, 195
- Roebber, E. 2019, *ApJ*, **876**, 55
- Roebber, E., & Holder, G. 2017, *ApJ*, **835**, 21
- Romani, R. W. 1989, in Timing Neutron Stars, ed. H. Ögelman & E. P. J. Heuvel, Vol. 262 (Berlin: Springer), 113
- Romano, J. D., Hazboun, J. S., Siemens, X., & Archibald, A. M. 2021, *PhRvD*, **103**, 063027
- Rosado, P. A., Sesana, A., & Gair, J. 2015, *MNRAS*, **451**, 2417
- Sampson, L., Cornish, N. J., & McWilliams, S. T. 2015, *PhRvD*, **91**, 084055
- Sardesai, S. C., & Vigeland, S. J. 2023, arXiv:2303.09615
- Sazhin, M. V. 1978, *SvA*, **22**, 36
- Sesana, A. 2013, *MNRAS*, **433**, L1
- Sesana, A., Haardt, F., Madau, P., & Volonteri, M. 2004, *ApJ*, **611**, 623
- Sesana, A., Vecchio, A., & Colacino, C. N. 2008, *MNRAS*, **390**, 192
- Siemens, X., Ellis, J., Jenet, F., & Romano, J. D. 2013, *CQGr*, **30**, 224015
- Speri, L., Porayko, N. K., Falxa, M., et al. 2023, *MNRAS*, **518**, 1802
- Stinebring, D. R., Ryba, M. F., Taylor, J. H., & Romani, R. W. 1990, *PhRvL*, **65**, 285
- Taylor, J. H., Fowler, L. A., & McCulloch, P. M. 1979, *Natur*, **277**, 437
- Taylor, S. R. 2021, Nanohertz Gravitational Wave Astronomy (Boca Raton, FL: CRC Press)
- Taylor, S. R., Baker, P. T., Hazboun, J. S., Simon, J. J., & Vigeland, S. J. 2018, enterprise_extensions, https://github.com/nanograv/enterprise_extensions
- Taylor, S. R., & Gair, J. R. 2013, *PhRvD*, **88**, 084001
- Taylor, S. R., Lentati, L., Babak, S., et al. 2017, *PhRvD*, **95**, 042002
- Taylor, S. R., Simon, J., Schult, L., Pol, N., & Lamb, W. G. 2022, *PhRvD*, **105**, 084049
- Taylor, S. R., van Haasteren, R., & Sesana, A. 2020, *PhRvD*, **102**, 084039
- Tiburzi, C., Hobbs, G., Kerr, M., et al. 2016, *MNRAS*, **455**, 4339

- Tomczak, A. R., Quadri, R. F., Tran, K.-V. H., et al. 2014, [ApJ](#), **783**, 85
- Vallisneri, M. 2020, libstempo: Python wrapper for Tempo2, Astrophysics Source Code Library, ascl:[2002.017](#)
- Vallisneri, M., Taylor, S. R., Simon, J., et al. 2020, [ApJ](#), **893**, 112
- van Haasteren, R., Levin, Y., McDonald, P., & Lu, T. 2009, [MNRAS](#), **395**, 1005
- van Haasteren, R., & Vallisneri, M. 2014, [PhRvD](#), **90**, 104012
- Vehtari, A., Gelman, A., Simpson, D., Carpenter, B., & Bürkner, P.-C. 2021, [BayAn](#), **16**, 667
- Vigeland, S. J., Islo, K., Taylor, S. R., & Ellis, J. A. 2018, [PhRvD](#), **98**, 044003
- Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, [NatMe](#), **17**, 261
- Wilcox, R. 2012, Introduction to Robust Estimation and Hypothesis Testing, Statistical Modeling and Decision Science (Amsterdam: Elsevier)
- Wyithe, J. S. B., & Loeb, A. 2003, [ApJ](#), **590**, 691
- Zic, A., Hobbs, G., Shannon, R. M., et al. 2022, [MNRAS](#), **516**, 410