The Astrophysical Journal, 859:47 (22pp), 2018 May 20 https://doi.org/10.3847/1538-4357/aabd3b

The NANOGrav 11 Year Data Set: Pulsar-timing Constraints on the Stochastic Gravitational-wave Background

Z. Arzoumanian[1], P. T. Baker[2,3], A. Brazier[4], S. Burke-Spolaor[2,3], S. J. Chamberlin[5], S. Chatterjee[4], B. Christy[6], J. M. Cordes[4], N. J. Cornish[7], F. Crawford[8], H. Thankful Cromartie[9], K. Crowter[10], M. DeCesar[11,38], P. B. Demorest[12], T. Dolch[13], J. A. Ellis[2,3,38,41], R. D. Ferdman[14], E. Ferrara[15], W. M. Folkner[16], E. Fonseca[17], N. Garver-Daniels[2,3], P. A. Gentile[2,3], R. Haas[18], J. S. Hazboun[19,20,38], E. A. Huerta[18,21], K. Islo[22], G. Jones[23], M. L. Jones[2,3], D. L. Kaplan[22], V. M. Kaspi[17], M. T. Lam[2,3,38], T. J. W. Lazio[16,24], L. Levin[25], A. N. Lommen[26], D. R. Lorimer[2,3], J. Luo[20], R. S. Lynch[27], D. R. Madison[28,39], M. A. McLaughlin[2,3], S. T. McWilliams[2,3], C. M. F. Mingarelli[29], C. Ng[10], D. J. Nice[11], R. S. Park[16], T. T. Pennucci[2,3,30,31,38], N. S. Pol[2,3], S. M. Ransom[9,28], P. S. Ray[32], A. Rasskazov[30,31,33], X. Siemens[22], J. Simon[16], R. Spiewak[22,34], I. H. Stairs[10], D. R. Stinebring[35], K. Stovall[12,38], J. Swiggum[22,38], S. R. Taylor[24,16,38], M. Vallisneri[16,24], R. van Haasteren[16,40], S. Vigeland[22,38], and W. W. Zhu[36,37]

(The NANOGrav Collaboration)

[1] Center for Research and Exploration in Space Science and Technology and X-Ray Astrophysics Laboratory, NASA Goddard Space Flight Center, Code 662, Greenbelt, MD 20771, USA
[2] Department of Physics and Astronomy, West Virginia University, P.O. Box 6315, Morgantown, WV 26506, USA
[3] Center for Gravitational Waves and Cosmology, West Virginia University, Chestnut Ridge Research Building, Morgantown, WV 26505, USA
[4] Department of Astronomy, Cornell University, Ithaca, NY 14853, USA
[5] Department of Astronomy and Astrophysics, Pennsylvania State University, University Park, PA 16802, USA
[6] Notre Dame of Maryland University, 4701 N. Charles Street, Baltimore, MD 21210, USA
[7] Department of Physics, Montana State University, Bozeman, MT 59717, USA
[8] Department of Physics and Astronomy, Franklin & Marshall College, P.O. Box 3003, Lancaster, PA 17604, USA
[9] University of Virginia, Department of Astronomy, P.O. Box 400325, Charlottesville, VA 22904, USA
[10] Department of Physics and Astronomy, University of British Columbia, 6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada
[11] Department of Physics, Lafayette College, Easton, PA 18042, USA
[12] National Radio Astronomy Observatory, 1003 Lopezville Rd., Socorro, NM 87801, USA
[13] Department of Physics, Hillsdale College, 33 E. College Street, Hillsdale, Michigan 49242, USA
[14] Department of Physics, University of East Anglia, Norwich, UK
[15] NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA
[16] Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA
[17] Department of Physics, McGill University, 3600 University St., Montreal, QC H3A 2T8, Canada
[18] NCSA, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA
[19] University of Washington Bothell, 18115 Campus Way NE, Bothell, WA 98011, USA
[20] Center for Advanced Radio Astronomy, University of Texas Rio Grande Valley, Brownsville, TX 78520, USA
[21] Department of Astronomy, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA
[22] Center for Gravitation, Cosmology and Astrophysics, Department of Physics, University of Wisconsin-Milwaukee, P.O. Box 413, Milwaukee, WI 53201, USA
[23] Department of Physics, Columbia University, New York, NY 10027, USA
[24] Theoretical AstroPhysics Including Relativity (TAPIR), MC 350-17, California Institute of Technology, Pasadena, California 91125, USA; srtaylor@caltech.edu
[25] Jodrell Bank Centre for Astrophysics, University of Manchester, Manchester, M13 9PL, UK
[26] Department of Physics and Astronomy, Haverford College, Haverford, PA 19041, USA
[27] Green Bank Observatory, P.O. Box 2, Green Bank, WV 24944, USA
[28] National Radio Astronomy Observatory, 520 Edgemont Road, Charlottesville, VA 22903, USA
[29] Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA
[30] Institute of Physics, Eötvös Loránd University, Pázmány P. s. 1/A, 1117 Budapest, Hungary
[31] Hungarian Academy of Sciences MTA-ELTE Extragalactic Astrophysics Research Group, 1117 Budapest, Hungary
[32] Naval Research Laboratory, Washington DC 20375, USA
[33] School of Physics and Astronomy and Center for Computational Relativity and Gravitation, Rochester Institute of Technology, Rochester, NY 14623, USA
[34] Centre for Astrophysics and Supercomputing, Swinburne University of Technology, P.O. Box 218, Hawthorn VIC 3122, Australia
[35] Department of Physics and Astronomy, Oberlin College, Oberlin, OH 44074, USA
[36] National Astronomical Observatories, Chinese Academy of Science, 20A Datun Road, Chaoyang District, Beijing 100012, People's Republic of China
[37] Max Planck Institute for Radio Astronomy, Auf dem Hügel 69, D-53121 Bonn, Germany
Received 2017 December 19; revised 2018 March 31; accepted 2018 April 8; published 2018 May 23

Abstract

We search for an isotropic stochastic gravitational-wave background (GWB) in the newly released 11 year data set from the North American Nanohertz Observatory for Gravitational Waves (NANOGrav). While we find no evidence for a GWB, we place constraints on a population of inspiraling supermassive black hole (SMBH) binaries, a network of decaying cosmic strings, and a primordial GWB. For the first time, we find that the GWB constraints are sensitive to the solar system ephemeris (SSE) model used and that SSE errors can mimic a GWB signal. We developed an approach that bridges systematic SSE differences, producing the first pulsar-timing array

[38] NANOGrav Physics Frontiers Center Postdoctoral Fellow.
[39] Jansky Fellow.
[40] Currently employed at Microsoft Corporation.
[41] Currently employed at Infinia ML 202 Rigsbee Avenue, Durham NC, 27701.

(PTA) constraints that are robust against SSE errors. We thus place a 95% upper limit on the GW-strain amplitude of $A_{\rm GWB} < 1.45 \times 10^{-15}$ at a frequency of $f = 1\ {\rm yr}^{-1}$ for a fiducial $f^{-2/3}$ power-law spectrum and with interpulsar correlations modeled. This is a factor of ∼2 improvement over the NANOGrav nine-year limit calculated using the same procedure. Previous PTA upper limits on the GWB (as well as their astrophysical and cosmological interpretations) will need revision in light of SSE systematic errors. We use our constraints to characterize the combined influence on the GWB of the stellar mass density in galactic cores, the eccentricity of SMBH binaries, and SMBH–galactic-bulge scaling relationships. We constrain the cosmic-string tension using recent simulations, yielding an SSE-marginalized 95% upper limit of $G\mu < 5.3 \times 10^{-11}$—a factor of ∼2 better than the published NANOGrav nine-year constraints. Our SSE-marginalized 95% upper limit on the energy density of a primordial GWB (for a radiation-dominated post-inflation universe) is $\Omega_{\rm GWB}(f)\ h^2 < 3.4 \times 10^{-10}$.

Key words: ephemerides – gravitational waves – inflation – methods: data analysis – pulsars: general – quasars: supermassive black holes

1. Introduction

Over the last two years, the gravitational-wave (GW) community has celebrated the first direct detection of GWs, generated by the coalescence of two ∼30 $M_\odot$ black holes (Abbott et al. 2016), as well as the first multimessenger observation of GWs with panspectral EM radiation, emitted during and after the final inspiral and merger of two neutron stars (Abbott et al. 2017). Pulsar-timing arrays (PTAs; Sazhin 1978; Detweiler 1979; Foster & Backer 1990) offer the opportunity of extending GW observations to the very-low-frequency spectrum (∼1–100 nHz). The discovery space here is populated by GWs from supermassive black hole binaries (SMBHBs) at galactic centers (see, e.g., Sesana et al. 2004; Sesana 2013), and possibly from more speculative sources of cosmological origin, such as cosmic strings (Damour & Vilenkin 2001; Ölmez et al. 2010) and/or a primordial GW background (GWB) produced by quantum fluctuations of the gravitational field in the early universe, amplified by inflation (e.g., Grishchuk 2005; Lasky et al. 2016).

The three major collaborations involved in this effort are the North American Nanohertz Observatory for Gravitational Waves (NANOGrav; McLaughlin 2013), the European Pulsar Timing Array (EPTA; Desvignes et al. 2016), and the Parkes Pulsar Timing Array (PPTA; Hobbs 2013). In addition, the International Pulsar Timing Array (IPTA; Verbiest et al. 2016) exists as an umbrella consortium for data sharing, coordinated timing campaigns, and joint GW analysis. The increasing sensitivity of PTAs is apparent in the ever-tightening upper limits (van Haasteren et al. 2011; Demorest et al. 2013; Shannon et al. 2013; Lentati et al. 2015; Shannon et al. 2015; Arzoumanian 2016) on the stochastic GWB from the unresolved superposition of SMBHB signals out to redshift $\lesssim$1.

The road toward detection lies not only through the accumulation of ever longer and more accurate time-of-arrival (TOA) data for larger arrays of monitored pulsars, but also through the development of powerful, robust, and reliable data analysis methods to demonstrate the presence of GWs in PTA data. In this article, we report substantial advances along both avenues. First, we present our stochastic-GW analysis of NANOGrav's largest and most sensitive data set so far, spanning 45 pulsars and 11.4 years. See Section 2 and Arzoumanian (2018, hereafter NG11) for more on this "11 year" data set. Second, we describe our statistical-inference framework, which was significantly augmented compared to our GW study of the nine-year data set (Arzoumanian 2016, hereafter NG9b). Improvements include a practical strategy to isolate the expected signature of stochastic GWs in our data—namely, the emergence of a long-timescale noise process that is common to all pulsars and the positive detection of interpulsar spatial correlations with a quadrupolar signature (Hellings & Downs 1983). This strategy is based on Bayesian model selection and is extensible to large arrays and data sets. Indeed, for the first time with a large pulsar array, we are able to report GW upper limits and GW versus noise ("detection") Bayes factors computed with likelihoods that include spatial correlations—such as the ones predicted by Hellings & Downs (1983)—a goal that had previously proved computationally infeasible beyond small arrays (Lentati et al. 2015).

This article also features a more robust, Bayesian–frequentist hybrid "optimal-statistic" analysis (Anholm et al. 2009; Demorest et al. 2013; Chamberlin et al. 2015), which complements our primary Bayesian approach. In addition, we employ a more flexible end-to-end approach for PTA GW searches to constrain astrophysical parameters (characterizing SMBHB populations and environments, as well as cosmic-string properties). This approach uses a set of GW-spectrum simulations that span the parameter-space region of interest and interpolates them by means of GPs (Williams & Rasmussen 2006; Taylor et al. 2017b), resulting in a flexible new model that is calibrated directly by detailed simulations.

Lastly, but perhaps most importantly, we report on how solar system ephemeris (SSE) errors can manifest as a false GWB signal in PTA data, for sufficiently long and high-quality data sets. The SSE is used to refer TOA measurements to an inertial frame located at the solar system barycenter (SSB). Previous GW searches treated the ephemeris as a fixed-parameter model without uncertainties. However, in the course of analyzing the 11 year data set, we discovered that adopting different ephemerides (among the last few published by the Jet Propulsion Laboratory (JPL); see Folkner et al. 2009, 2014, 2016; Folkner & Park 2016) leads to significantly different upper-limit and model-comparison statistics. As PTA data sets become larger, longer, and more precise, our GW searches will continue to uncover systematic effects that will limit our sensitivity unless handled appropriately. To this end, we have developed a physical model of ephemeris uncertainties, and we demonstrate that it makes our analysis insensitive to the choice among recent ephemerides.

This paper is laid out as follows: methodological advances are discussed in Section 3. In Sections 4 and 5, we report GW upper limits and detection Bayes factors based on the 11 year data set, as well as new constraints on astrophysical and cosmological sources of low-frequency GWs. In Section 6, we present our conclusions and discuss prospects for future observations.

For the busy reader, the following summarizes the most consequential results:

1. Once we take ephemeris uncertainty into account, we find Bayesian model comparison to be inconclusive on the presence of a GWB-like signal in the data (with signal-versus-noise and spatial-correlation Bayes factors both $\sim$1). Adopting one of the fixed JPL ephemerides leads to signal-versus-noise Bayes factors as high as 26 ± 2 in favor of a GWB-like signal (for JPL ephemeris DE430), suggesting that systematic ephemeris errors can masquerade as GWs—and conversely that modeling these errors can subtract power from a putative GWB signal. This degeneracy will be resolved over the next few years as we collect longer and larger data sets, and as ephemeris accuracy improves with data from current NASA missions.
2. Accounting for ephemeris uncertainty, the 95% Bayesian upper limit on a fiducial $f^{-2/3}$ GW spectrum from SMBHBs is $A_{\rm GW}^{95\%} = 1.45(2) \times 10^{-15}$ at $f = 1\,{\rm yr}^{-1}$ (when modeling spatial correlations; $1.34(1) \times 10^{-15}$ when omitting them). This value is modestly improved from the nine-year result of $A_{\rm GW}^{95\%} = 1.5 \times 10^{-15}$, which omitted correlations, and assumed JPL ephemeris DE421 as a fixed-parameter model without uncertainties. Note, however, that reprocessing the nine-year data set accounting for ephemeris uncertainties leads to $A_{\rm GW}^{95\%} = 2.91(2) \times 10^{-15}$ (when modeling spatial correlations; $2.67(2) \times 10^{-15}$ omitting them). We expect that recently published limits from other PTAs (such as Shannon et al. 2015) would be likewise revised upwards. Our 11 year upper limits assuming individual fixed-model ephemerides range from 1.53 to 1.78×10^{-15} (when modeling spatial correlations), again suggesting that ephemeris errors can mimic GW-like signals.
3. We place the first joint constraints on the galaxy properties and binary evolution parameters with the greatest impact on the spectral shape and amplitude of the GWB from SMBHBs. Previous work, such as that undertaken in NG9b, has always utilized an amplitude or spectral shape assumption before inferring any astrophysical constraints. This improved methodology allows for the first robust PTA limits on the $M_{\rm BH}$–$M_{\rm bulge}$ relation and shows that the NANOGrav 11 year data set prefers a relation that is lower (in terms of the relation's y-intercept) than that reported in Kormendy & Ho (2013).
4. Using a model of cosmic-string-generated GW spectra that interpolates among extensive string-network simulations (Blanco-Pillado & Olum 2017), we place a 95% upper limit of $5.3(2) \times 10^{-11}$ on the string tension $G\mu/c^2$ for a reconnection probability $p = 1$. This result is marginalized over ephemeris uncertainties, but ignores interpulsar spatial correlations. (Including these is still too taxing computationally; however, we argue that our upper limits assuming a variable power-law exponent, described in Section 4.1, are affected modestly by correlations, and so should be the cosmic-string result.) Previous studies reported limits of 1.3×10^{-10} (NG9b) and 8.6×10^{-10} (Lentati et al. 2015), although different prior assumptions and the lack of ephemeris modeling preclude a direct comparison.
5. Lastly, we can interpret power-law GWBs with different fixed exponents as a primordial background amplified through inflation, with post-inflationary eras characterized by different equations of state. Assuming a radiation-dominated post-inflation universe and a tensor index $n_t = 0$ (corresponding to a scale-invariant spectrum) leads to a 95% upper limit of $3.4(1) \times 10^{-10}$ on the GW energy density $\Omega_{\rm GW}(f)h^2$ at $f = 1\,{\rm yr}^{-1}$, again marginalizing over ephemeris uncertainty but ignoring interpulsar spatial correlations.

2. The 11 Year Data Set

Our analyses throughout this paper make use of the NANOGrav 11 year data set, which consists of the TOA data and pulsar-timing models recently presented in NG11 and is publicly available online.[42] This data set is derived from timing observations of 45 millisecond pulsars between 2004 July 30 to 2015 December 31. The first five years of data on 17 pulsars constituted the NANOGrav five-year data set, which we previously published in NG5. The five-year data set was augmented by four years of data, reported as the nine-year data set in Arzoumanian et al. (2015, hereafter NG9a), which came with the substantial improvements of the new broadband instrumentation, a nearly twofold increase in the timing baseline for the original 17 pulsars, and a more than twofold increase in the total number of observed sources to 37 pulsars. The present extension of the nine-year data set is composed of two years of data that were observed and processed in a nearly identical fashion to the previous augmentation, with the addition of nine pulsars and the removal of one (see NG11 for full details). Here we briefly review the instrumentation, observations, and basic data reduction of the entire data set, referring the reader to NG11, NG9a, and references therein for a thorough description. A sky map of the pulsars in this data set is shown in Figure 1, with indicators of the time span and data volume for each pulsar.

2.1. Observations

We obtained all data using the 100 m Robert C. Byrd Green Bank Telescope (GBT) of the Green Bank Observatory[43] and the 305 m William E. Gordon Telescope (Arecibo) of Arecibo Observatory.[44] Sources within Arecibo's declination range ($0° < \delta < 39°$) were observed there due to its superior sensitivity, and only pulsars J1713+0747 and B1937+21 were observed at both telescopes. Excluding early portions from the five-year data set, we observed each source roughly once a month for the entire data set. In addition, some pulsars have been observed weekly in a campaign to increase our sensitivity to individual sources of GWs (Arzoumanian et al. 2014, Section 6.1). Specifically, two pulsars have been observed weekly at the GBT since 2013 (PSRs J1737+0747 and J1909 −3744), and five pulsars have been observed weekly at Arecibo since 2015 (PSRs J0030+0451, J1640+2224, J1713 +0747, J2043+1711, and J2317+1439).

During most epochs,[45] we observed sources in two widely separated frequency bands in order to accurately remove the frequency-dependent dispersion delay introduced by the ionized interstellar medium (ISM). At the GBT, we used

[42] data.nanograv.org
[43] greenbankobservatory.org/telescopes/gbt/
[44] http://outreach.naic.edu/ao/
[45] This excludes the weekly observations, which were performed at 1.4 GHz only, as well as epochs for which receivers were unavailable for technical reasons.

the 820 MHz and 1.4 GHz receivers for all observations. Since mechanical and time constraints prohibit alternating continually between the two receivers, observations in the two bands were always separated, typically by several days. At Arecibo, we observed all pulsars at 1.4 GHz, plus a second frequency band (centered at either 430 MHz or 2.3 GHz) chosen depending on the spectrum and ISM characteristics of each pulsar.[46] Pulsars observed at Arecibo are always observed in the two frequency bands one after another, separated by a few minutes.

For approximately the first six years, data were acquired with an identical pair of backend instruments that have since been decommissioned (GASP at the GBT, ASP at Arecibo). Since 2010 and 2012, respectively, the broadband-capable backend clones GUPPI (at the GBT) and PUPPI (at Arecibo) have been used for taking data.

2.2. Processing and TOA Data

The raw data products are folded light curves (average, uncalibrated flux density as a function of rotational phase, divided into 2048 phase bins) as a function of time, radio frequency, and polarization. These data were cleaned of radio-frequency interference in several steps, polarization-calibrated according to standard techniques, and averaged in time and frequency. The final time resolution was either 30 minutes or 2.5% of the binary period, whichever was shorter (approximately two-thirds of our pulsars are in binary systems). This length of time corresponds to how long we point our radio telescopes at each pulsar during a single observation session, where we fold many individual pulses then convolve with a pulse-profile template to compute a single TOA. This is necessary to achieve high timing precisions of ∼100 ns. The final frequency resolution varied between 1.5 and 12.5 MHz, depending on the receiver–backend combination.

The five-year TOA data set was left mostly untouched as a subset of the 11 year data set, except for reprocessing under DE436. All of the GUPPI and PUPPI profile data, however, were reprocessed from scratch to make a consistent set of TOAs. The TOAs were generated using standard template-matching cross-correlation methods, using only the total intensity profiles, producing one TOA per frequency channel per temporal subintegration. Existing template profiles were reused for pulsars that were part of the nine-year data set and created for new pulsars.

An additional set of procedures culled "outlier," low signal-to-noise (non-Gaussian distributed), or otherwise corrupt TOAs from the data set using methods described in Vallisneri & van Haasteren (2017). The 11 year data set comprises a total of 309,201 TOAs. All data reduction was completed using PSRCHIVE.[47] (Hotan et al. 2004) and custom NANOGrav processing scripts.[48]

2.3. Timing Models and Noise Analysis

Timing models from the nine-year data set were refit to the extended data set and updated to include new parameters when deemed necessary on the basis of statistical significance tests. We fit timing models for newly added pulsars using a procedure similar to that described in NG9a. All timing models were created or updated using the standard timing software TEMPO[49] and TEMPO2[50] (Edwards et al. 2006; Hobbs et al. 2006), and cross-checked for consistency.

A standard noise model was also fit simultaneously with the timing model as described in NG9a and NG11. Each pulsar's white-noise model includes a scale parameter on the TOA uncertainties (EFAC), an added variance (EQUAD), and a per-epoch variance (ECORR) for each observing system (i.e., a unique combination of backend and receiver). In addition, a red-noise process for each pulsar was modeled by a power-law spectral density described by an amplitude and spectral index. The inclusion of a red process in the noise model was not favored by all pulsars, but we include it in all subsequent analyses since this does not affect parameter constraints. In the analyses described in the subsequent sections, we vary the pulsars' red-noise parameters and the parameters of the GWB, but fix the white-noise parameters. Allowing the white-noise parameters to vary does not alter the results, but it significantly increases the computation time.

The SSE model used for the original analysis of the five-year data set (Demorest et al. 2013) was DE405 (Standish 2004), while for the nine-year data set (Arzoumanian et al. 2015) all data (whether new or from the five-year data set) was modeled with DE421 (Folkner et al. 2009). For the 11 year data set, we use DE436 (Folkner & Park 2016) as the fiducial SSE under which the data is processed and released. We do not need separate data set releases for the different SSEs that we investigate in the following, since our GWB analysis incorporates marginalization over all affected processes, such as the individual timing and red-noise models.

3. Data Analysis Methods

Characterizing all deterministic and noise processes in each pulsar, as well as teasing out a putative GWB signature from the cross-correlation of large data sets, requires a robust and sophisticated statistical framework. In the following we describe the major new features of the NANOGrav PTA analysis framework, as updated from NG9b. Section 3.1 describes our use of Bayesian inference as it pertains to computing GWB upper limits and detection statistics. Section 3.2 outlines how a GWB manifests in our data as a long-timescale stochastic process with a distinctive correlation signature between pulsars. In Section 3.3, we describe how the SSE model appears in our PTA and our new Bayesian scheme to mitigate its uncertainties. The structure of our generative signal and noise model is outlined in Section 3.4, followed in Section 3.5 by the definition of our frequentist estimator for the GWB amplitude and significance. Finally, in Section 3.6, we list and provide links for all open-source software used in our GWB analysis.

3.1. Bayesian Methods

We primarily employ Bayesian inference (see, e.g., Gregory 2005) to extract physical information from our data, deriving marginalized posterior distributions and credible regions, basing upper limits on credible intervals, and relying on ratios of evidences (a.k.a. Bayes factors) to compare models with different assumptions and parametrizations. We explore our high-dimensional parameter space stochastically using the parallel-tempering Markov Chain Monte Carlo (MCMC)

[46] Pulsar J2317+1439 was originally observed with the 327 and 430 MHz receivers, but in 2014 we replaced the former with the 1.4 GHz receiver.
[47] psrchive.sourceforge.net
[48] github.com/demorest/nanopipe
[49] tempo.sourceforge.net
[50] bitbucket.org/psrsoft/tempo2.git

sampler (Ellis & van Haasteren 2017b) described in the appendices of Arzoumanian et al. (2014).

To place upper limits on the GWB amplitude $A_{\rm GWB}$, we compute its posterior density distribution $p(A_{\rm GWB}|\mathcal{D})$ (with $\mathcal{D}$ the data) by giving $A_{\rm GWB}$ a uniform prior distribution that encloses the support of the likelihood, and we estimate the 95% quantile by means of the empirical cumulative-distribution function estimator $\hat{A}^{95\%}_{\rm GWB}$ (Wilcox 2016). We approximate the standard error of $\hat{A}^{95\%}_{\rm GWB}$ as

$$\frac{\sqrt{x(1-x)/N}}{p(A_{\rm GWB}=\hat{A}^{95\%}_{\rm GWB}|\mathcal{D})}, \quad (1)$$

with $x = 0.95$ and N the number of (quasi-)independent samples[51] in the chain.

As our PTA data set becomes longer and more sensitive, we expect that evidence for the presence of GWs will emerge in two phases: first as red-spectrum processes with the same amplitude in each pulsar and with spectral slope consistent with an SMBHB population; later (perhaps several years), and conclusively, as Hellings–Downs spatial correlations predicted for an isotropic GWB. We note that anisotropic GWBs will have different (but predictable) spatial correlations (Mingarelli et al. 2013; Taylor & Gair 2013; Gair et al. 2014; Mingarelli & Sidery 2014).

Correspondingly, we characterize evidence for a GWB in the 11 year data set in two steps. We first obtain the Bayes factor for a data set model that includes a red-spectrum process with common statistical properties in all pulsars (but is uncorrelated between them) against a model with only per-pulsar noise processes. This is signal-versus-noise model selection. We then obtain the Bayes factor for Hellings–Downs interpulsar spatial correlations versus no correlations at all. This is spatial-correlation model selection, which we consider the definitive scheme for GWB detection. We also perform variants of these comparisons—for instance, we compare the Hellings–Downs and uncorrelated process against processes with monopolar (akin to long-timescale clock errors) and dipolar (akin to SSE errors) spatial correlations.

In all cases, we adopt bounded log-uniform priors for $A_{\rm GWB}$ and all other red-process amplitudes. We adopt two different techniques to compute Bayes factors according to the relation between the models in the comparison.

For nested models (in our case, a signal-plus-noise model $\mathcal{H}_1$ and a noise-only model $\mathcal{H}_0$ obtained by fixing the GW amplitude to 0), we employ the Savage–Dickey formula (Dickey 1971)

$$\mathcal{B}_{10} \equiv \frac{\text{evidence}[\mathcal{H}_1]}{\text{evidence}[\mathcal{H}_0]} = \frac{p(A_{\rm GWB}=0|\mathcal{H}_1)}{p(A_{\rm GWB}=0|\mathcal{D},\mathcal{H}_1)}, \quad (2)$$

where the numerator and denominator are, respectively, the prior and posterior probability density of $A_{\rm GWB} = 0$ in the embedding model $\mathcal{H}_1$. We generate a sampling of $p(A_{\rm GWB}|\mathcal{D},\mathcal{H}_1)$ via MCMC, and we approximate $p(A_{\rm GWB} = 0|\mathcal{D},\mathcal{H}_1)$ as the normalized fraction of samples in the lowest-amplitude bin, averaging the estimate over a range of bin sizes. The standard error of this average yields an estimate of uncertainty for the Bayes factor.

[51] Quasi-independence here refers to samples separated by one autocorrelation chain length.

For disjoint models (in our case, a model consisting of a Hellings–Downs-correlated red process plus pulsar noise versus a model consisting of a common-amplitude, spatially uncorrelated red process plus pulsar noise) we use a product-space method (Carlin & Chib 1995; Godsill 2001; Hee et al. 2016). In this method, we define a super-model that contains all parameters from all models under consideration, as well as an additional model-indexing variable that determines which model is "active" and used to evaluate the likelihood.[52] (In our example, where the parameters are actually the same in both models, the index variable would simply toggle Hellings–Downs correlations in the evaluation of the likelihood.) The ratios of posterior probabilities for two model indices approximate the corresponding Bayes factor. We follow Cornish & Littenberg (2015) to estimate Bayes-factor uncertainties.

Evaluating the multipulsar likelihood is very computationally expensive when we account for interpulsar spatial correlations. In that case, we accelerate inference by running at least 10 parallel copies of each spatially correlated analysis. These subchains can then be concatenated to form a much larger chain. Each subchain is analyzed to determine that it has "burned in"[53] before combining it with others. To derive upper limits and Savage–Dickey Bayes factors, we simply append the subchains together and proceed as described above. For product-space Bayes factors, we obtain the factor itself from the combined subchains, but we estimate uncertainties in each subchain separately, then add them in quadrature (Cornish & Littenberg 2015).

Arbitrary rules of thumb have been given to interpret the statistical significance of Bayes factors of different magnitudes (see, e.g., Jeffreys 1961; Kass & Raftery 1995), but it is hard to find agreement beyond the trivial statement that factors $\sim$1 are inconclusive, while very large or small factors point to a strong preference for either model. In the context of a detection scheme, it seems appropriate to examine the frequentist distribution of the Bayes factor and to set detection thresholds as a function of false-alarm probability (Vallisneri 2012). The sky-scramble and phase-shift methods (Cornish & Sampson 2016; Taylor et al. 2017a) have been proposed to produce a background distribution of the Bayes factor for which spatial correlations are effectively removed from the data. By contrast, we currently lack a practical approach to establish the significance of a common uncorrelated process; such an approach would likely involve a combination of inference runs on simulated data and cross-validation experiments, such as comparing results for subsets of the data set. As we shall see, all the ephemeris-marginalized Bayes factors obtained in this paper are close to unity and can be deemed inconclusive without a frequentist analysis.

3.2. Gravitational-wave Strain Spectrum

The observed timing residuals due to a GWB with characteristic strain $h_c(f)$ are described by the cross-power

[52] This variable is technically discrete, but it can be sampled continuously and cast to an integer to choose the active model.

[53] In MCMC analysis, some early sampled points must be disregarded before the chain can be considered to be sampling from the true posterior probability distribution. The disregarded early portion of the chain is called the "burn in" stage.

spectral density

$$S_{ab}(f) = \Gamma_{ab}(f)\frac{{h_c}^2(f)}{12\pi^2 f^3}, \quad (3)$$

where Γ_{ab} is the overlap reduction function (ORF), which describes correlations between pulsars a and b in the array. In the case of an isotropic background from SMBHBs, the ORF is given by Hellings & Downs (1983; hereafter referred to as H.–D. correlations). Other correlated effects such as systematic errors in the SSE or clocks can also be described by a timing-residual spectrum that includes a different ORF.

In this paper, we consider four models of the GWB spectrum:

Power-law spectrum: A population of inspiraling SMBHBs in circular orbits, evolving by GW emission alone, produces a characteristic GW-strain spectrum, expressed as

$$h_c(f) = A_{\rm GWB}\left(\frac{f}{\rm yr^{-1}}\right)^{\alpha} \quad (4)$$

with $\alpha = -2/3$ (Phinney 2001). Different spectral slopes can be used to model relic radiation from the early universe, under different assumptions for the equation of state of the universe post-inflation/pre-big bang nucleosynthesis (see Section 5.3). We find it expedient to perform our analysis in terms of the timing-residual spectral index $\gamma = 3 - 2\alpha$, such that

$$S_{ab}(f) = \Gamma_{ab}\frac{{A_{\rm GWB}}^2}{12\pi^2}\left(\frac{f}{\rm yr^{-1}}\right)^{-\gamma}{\rm yr}^3. \quad (5)$$

The fiducial SMBHB $\alpha = -2/3$ then corresponds to $\gamma = 13/3$.

Broken power-law spectrum: If SMBHBs remain coupled to the dynamics of their galactic environments as they evolve into the nanohertz band, the nanohertz GW-strain spectrum will be more complex than described by Equation (4). This may be the case if three-body scattering of stars from the galactic-center loss cone (e.g., Quinlan 1996; Sesana et al. 2006) or interaction with a viscous circumbinary disk (e.g., Haiman et al. 2009; Kocsis & Sesana 2011) is a stronger dynamical influence than GW emission at wide orbital separations. When the binary reaches milliparsec separations, GW emission will always be dominant. Sampson et al. (2015) introduced a broken power-law model,

$$h_c(f) = A_{\rm GWB}\frac{(f/{\rm yr}^{-1})^{\alpha}}{(1 + (f_{\rm bend}/f)^{\kappa})^{1/2}}, \quad (6)$$

to model such spectra, where the slope transitions from positive at low frequencies to the canonical $-2/3$ at higher frequencies. The frequency at which the transition occurs encodes information about the typical binary's orbital evolution and astrophysical environment.

Free spectrum: To characterize the GW-strain sensitivity of our data set as a function of frequency, we adopt independent uniform priors for the dimensionless strain amplitudes of each sine–cosine pair of red-process Fourier components (see Section 3.4), corresponding to frequencies k/T, with $k = 1, \ldots, N$, where T is the longest time span in the combined data set and N (set to 50 in this paper) is the number of Fourier component pairs. We then derive a joint posterior for all amplitudes.

GP spectrum emulation: This model was introduced by Taylor et al. (2017b) as a way to perform searches that are directly informed by detailed source-astrophysics simulations and to sample the posteriors of the binary environment and dynamics parameters that affect the GW spectrum without generating a new simulation for each likelihood evaluation. In practice, we perform simulations over a grid in the parameter space of interest, and for each simulation we compute the GW characteristic strain spectrum. We then train a GP (Williams & Rasmussen 2006) to interpolate over all spectra in parameter space, allowing spectral amplitudes to be predicted at any other point with an associated normal uncertainty. We then use these predictions and uncertainties as priors on the strain amplitude at each frequency within the free-spectrum model.

3.3. SSE Errors and Uncertainties

An SSE is used in pulsar timing to convert observatory TOAs to an inertial frame centered at the solar system barycenter, factoring out all effects due to Earth's motion. The dominant correction to the TOAs is the Roemer delay—the classical light-travel time between the geocenter and the solar system barycenter. Pulsar-timing studies have typically relied on the latest SSE released by JPL, adopting it as a model with fixed parameters—that is, without including any SSE parameter uncertainties or corrections in timing-model fits. In the early stages of our analysis of the NANOGrav 11 year data set, we became aware that the choice of SSE among the latest few released by JPL has a measurable impact on our GWB upper limits and model-comparison Bayes factors. Indeed, the abundance and precision of NANOGrav's measurements are now such that the accuracy to which we can estimate the Earth's orbit around the SSB limits our sensitivity to GWs. SSE errors have been speculated as a source of potential bias in PTA GW detection efforts (Tiburzi et al. 2016), but this paper marks the first time that this effect has been rigorously studied with real data sets.

The JPL SSEs,[54] as well as the French INPOP,[55] fit the orbits and masses of a large set of solar system bodies to a heterogeneous data set collected over the last few decades, using spacecraft ranging, direct planetary radar ranging, spacecraft VLBI, and (for the Moon) laser ranging of retroreflectors left by the *Apollo* missions. The orbits are integrated numerically from initial conditions ("epoch" positions and velocities), which are the parameters that are fit for, together with other quantities such as the masses of minor solar system bodies (but not planet masses, which are estimated separately from observed motions in planetary systems; Folkner et al. 2009). The resulting SSEs are distributed as Chebyshev polynomials over a range of dates; notably, they do not include estimates of orbit uncertainties and of possible systematics.

To investigate the effects of SSE errors, we repeated all upper-limit and model-comparison analyses in this paper using the four most recent JPL SSEs: DE421, released in 2008 (Folkner et al. 2009); DE430 (Folkner et al. 2014); DE435

[54] https://ssd.jpl.nasa.gov/?ephemerides
[55] https://www.imcce.fr/inpop

(Folkner et al. 2016); and DE436 (Folkner & Park 2016)); for the simplest analysis, we used also the French INPOP13c (Fienga et al. 2014). The orbit of Earth relative to the Sun is consistent at the 10 m level across these ephemerides, after accounting for an overall rotation w.r.t. the International Celestial Reference Frame, which originates from updated very-long-baseline-interferometry observations of spacecraft at Mars. However, the orbit of the Sun w.r.t. the SSB and (therefore) the orbit of the Earth w.r.t. the SSB match only at the 100 m level. This discrepancy is attributed to differences in the estimated masses and positions of Jupiter, Uranus, and Neptune. Hence, our GW analysis shows significant systematic differences among the upper limits and Bayes factors computed using different ephemerides. Near-future efforts may lead to improvements in the ephemeris accuracy that are appropriate for pulsar timing, namely (i) estimates of Jupiter's orbit will be improved by including *Juno* spacecraft data in the SSE fit, (ii) ranging data from *Cassini* may better estimate the mass of Uranus, (iii) *Gaia* data may improve orbit estimates for Uranus and Neptune, and finally, (iv) pulsar-timing data may be used to improve the estimate of Neptune's mass.

We account for the differences between SSEs by developing a physical model (BayesEphem) that corrects Earth's tabulated orbit using 11 parameters. Four of these correspond to perturbations in the masses of the outer planets and generate corrections $-(\delta M_i/M_{\rm tot})\,\boldsymbol{r}_i(t)$, where δM_i is the outer planet's mass correction, $M_{\rm tot}$ is the total mass of the Solar System, and $\boldsymbol{r}_i(t)$ is the outer planet's orbit. One parameter describes a rotation rate about the ecliptic pole: this accounts for differences in the estimated semimajor axis of the Earth–Moon-barycenter orbit, which gives rise to a linear rate in estimated ecliptic longitude. Since the orbit of the Sun about the SSB is largely influenced by Jupiter, and since the Jovian period is comparable to the span of our data set, we also include corrections to Earth's orbit generated by perturbing Jupiter's average orbital elements. These corrections have the form $-(M_J/M_{\rm tot})\,(\partial \boldsymbol{r}_J(t)/\partial a_J^{\mu})\,\delta a_J^{\mu}$, where the partial derivatives encode the changes in Jupiter's orbit as we change the orbital elements, and where the six δa_J^{μ} are the orbital-element perturbations (which we define using Brouwer & Clemence's 1961 "set-III" parameters). By contrast, Saturn's orbit is constrained more strongly by available data, while Uranus and Neptune have large orbit uncertainties but very long periods, so they can only generate linear-in-time Roemer biases that are absorbed by fitting pulsar periods.

Thus, we present GW upper limits and model-comparison Bayes factors that are marginalized over these SSE uncertainty parameters. We regard these BayesEphem limits and Bayes factors as our fiducial results in this paper. To derive them, we constrain the outer-planet masses using the current IAU best estimates (Jacobson et al. 2000, 2006; Jacobson 2009, 2014; IAU 2017) and use IAU uncertainties to set Gaussian priors. The rate of rotation about the ecliptic pole is left unconstrained. We experimented with setting priors for Jupiter's orbital elements using estimated uncertainties,[56] but we find better results using uninformative priors. This is not surprising, because Jupiter's orbital elements are highly correlated with those of the other planets, and our linearized correction of Jupiter's orbit cannot account for those correlations. Nevertheless, the resulting variations of Earth's orbit are comparable with the systematic differences that we observe across JPL SSEs, which we take as evidence that the BayesEphem uncertainty parameters are representative of true SSE uncertainties.

Our Bayesian-inference studies produce orbital-element posteriors for Jupiter corresponding to position offsets at the level of $\sim$100 km. We defer the full details of our investigations of SSE uncertainties and systematics to an upcoming paper, where we compare reconstructed Jupiter orbits from our analysis to those from the JPL ephemerides and discuss potential modeling improvements to BayesEphem.

3.4. Data Model and Likelihood

Except for GP spectrum emulation and for the treatment of SSE errors, the data model used in this paper matches that of NG9b closely, so we refer the reader to that publication for an overview of noise modeling, marginalization over timing-model parameters, our rank-reduced formalism for time-correlated processes (e.g., timing noise or GWB), and the PTA likelihood.

The rank-reduced formalism refers to the expansion of processes on a sine–cosine Fourier basis with frequencies k/T, where T is the span between the minimum and maximum TOAs in the array. The number of basis vectors is chosen to be high enough that inference results are insensitive to adding more: we use 30 for all applications except for the free-spectrum GWB model, for which we use 50.

As for the PTA likelihood, we introduced a significant change compared to NG9b. "ECORR" (jitter-like) noise is fully correlated for simultaneous observations at different observing frequencies, but fully uncorrelated in time. In NG9b, we treated ECORR degrees of freedom by assigning them "exploder" basis vectors and then analytically marginalizing their coefficients simultaneously with timing-model, red-noise, and GWB coefficients. Doing so becomes computationally prohibitive when H.–D. correlations are included. In this paper, we include ECORR noise as block-diagonal entries (one block per epoch per backend–receiver system) in the otherwise diagonal white-noise covariance matrix and invert the matrix using the fast Sherman & Morrison (1950) formula. Doing so eliminates a significant computational bottleneck.

As in NG9b, computational efficiency is also helped by fixing all white-noise parameters to their 1D maximum a posteriori values from single-pulsar noise studies. This choice is justified empirically by the very small variance of white-noise parameters.

Our upper-limit and model-comparison studies are performed under a variety of assumptions about the presence of red-spectrum processes: in addition to individual red-spectrum timing noise for each pulsar, we model the GWB as a spatially uncorrelated common process (a computational simplification appropriate in the weak-GWB limit, used in NG9b) and as an H.–D.-correlated common process (our fiducial GWB model); we also consider common processes with different correlations (dipolar, as appropriate for SSE errors, and monopolar, as appropriate for long-timescale clock errors). Table 1 describes the nine models used in this paper, which are labeled 1, 2A–D, and 3A–D. In model-class 1, only intrinsic pulsar noise processes are included; in model-class 2, there are intrinsic pulsar noise processes as well as non-GW noise processes that induce interpulsar spatial correlations (such as clock and SSE errors); in model-class 3, we include a GWB signal. The roman

[56] W. M. Folkner & R. S. Park (2017, private communication) estimate uncertainties in Jupiter and Saturn orbits by comparing fits that use independent subsets of the data for each planet.

Table 1
Spatially Correlated Red-noise Processes Used in Our Analysis

	Model								
Red-noise Process	1	2A	2B	2C	2D	3A	3B	3C	3D
Intrinsic (per pulsar)	✓	✓	✓	✓	✓	✓	✓	✓	✓
Uncorr. common		✓							
H.-D. corr. common						✓	✓	✓	✓
Dipole corr. common			✓	✓			✓	✓	
Monopole corr. common				✓	✓			✓	✓

Note. All models include intrinsic white-noise and red-noise processes in each pulsar; additional common processes (with the same characteristic amplitude and spectrum in every pulsar) can be uncorrelated or have Hellings–Downs (GW-like), dipolar (ephemeris-error-like), and monopolar (clock-error-like) spatial correlations. Model 2A (uncorrelated common process) was used to derive the main results of NG9b; model 3A (Helling–Downs-correlated common process) is the fiducial model used to constrain the GWB in this publication.

characters given after the model-class number indicate the specific combination of noise and signal processes forming the model.

We perform each analysis by adopting each of the DE421, DE430, DE435, and DE436 (and occasionally INPOP13c) ephemerides as fixed-parameter models and by marginalizing over SSE uncertainties using BayesEphem. Our Bayesian priors for all parameters are described in Table 2.

3.5. Optimal Statistic

As in NG9b, we perform a frequentist GWB analysis using the optimal statistic $\hat{A}^2_{\rm GWB}$, a point estimator for the amplitude of an isotropic GW stochastic background (Anholm et al. 2009; Chamberlin et al. 2015). This statistic accounts implicitly for interpulsar spatial correlations. The estimator is derived by maximizing the PTA likelihood analytically, and it can be written as

$$\hat{A}^2_{\rm GWB} = \frac{\sum_{ab} \delta \boldsymbol{t}_a^T \boldsymbol{P}_a^{-1} \tilde{\boldsymbol{S}}_{ab} \boldsymbol{P}_b^{-1} \delta \boldsymbol{t}_b}{\sum_{ab} {\rm Tr}\,(\boldsymbol{P}_a^{-1} \tilde{\boldsymbol{S}}_{ab} \boldsymbol{P}_b^{-1} \tilde{\boldsymbol{S}}_{ba})}, \tag{7}$$

where $\delta \boldsymbol{t}_a$ is the vector of timing residuals for pulsar a, $\boldsymbol{P}_a = \langle \delta \boldsymbol{t}_a \delta \boldsymbol{t}_a^T \rangle$ is the autocovariance matrix of the residuals, and $\hat{A}^2_{\rm gw} \tilde{\boldsymbol{S}}_{ab} = \boldsymbol{S}_{ab} = \langle \delta \boldsymbol{t}_a \delta \boldsymbol{t}_b^T \rangle|_{a \neq b}$ is the cross-covariance matrix between the residuals for pulsars a and b. The average signal-to-noise ratio (S/N) of the optimal statistic is

$$\langle \rho \rangle = A^2_{\rm gw} \left[\sum_{ab} {\rm Tr}\,(\boldsymbol{P}_a^{-1} \tilde{\boldsymbol{S}}_{ab} \boldsymbol{P}_b^{-1} \tilde{\boldsymbol{S}}_{ba}) \right]^{1/2}, \tag{8}$$

which is a measure of the significance of interpulsar spatial correlations. When drawing comparisons between results produced using this frequentist technique and our Bayesian techniques, the relevant model selection is between models 3A and 2A.

We use two procedures to compute $\hat{A}^2_{\rm GWB}$. In the more conventional fixed-noise analysis, we compute $\hat{A}^2_{\rm GWB}$ at fixed values of the pulsar red-noise parameters $A_{\rm red}$ and $\gamma_{\rm red}$. The red-noise parameters are the values that jointly maximize the likelihood, as found in a Bayesian parameter-estimation study that includes pulsar red-noise and a common red-noise process. In the newer noise-marginalized analysis (S. J. Vigeland et al. 2017, in preparation), we use posterior samples from a Bayesian study to marginalize the optimal statistic over pulsar red-noise parameters. This results in distributions for both $\hat{A}^2_{\rm GWB}$ and the S/N, rather than a single value of $\hat{A}^2_{\rm GWB}$ and a corresponding S/N. In both cases, pulsar white-noise parameters are fixed to their maximum-likelihood values, as determined individually for each pulsar with Bayesian inference. As discussed in S. J. Vigeland et al. (2017, in preparation), simulations show that the noise-marginalized technique produces more accurate estimates of $A_{\rm GWB}$ compared to the fixed-noise technique. This is because the pulsar red-noise parameters are highly covariant with common-process red-noise parameters, so the fixed-noise analysis tends to systematically underestimate the amplitude and significance of common signals.

3.6. Software

We generated most of the results in this paper using the open-source software package `NX01`[57], which implements the PTA likelihood and priors. `NX01` was validated on a wide range of problems, including several 11 year analyses, by cross-comparison with the well-established `PAL2`[58] (Ellis & van Haasteren 2017a) and with NANOGrav's new flagship package, `enterprise`[59]. We perform MCMC using `PTMCMCSampler`[60] (Ellis & van Haasteren 2017b), which implements a variety of proposal schemes (adaptive Metropolis, differential evolution, parallel tempering, etc.), which can be used together in the same run.

As a companion to this paper, we are releasing a `Docker`[61] image that contains a full stack of our software (including all required libraries) and that can be used to reproduce the upper limits, Bayes factors, as well as many of the figures of this paper, using `enterprise`.

4. Results

All results in this paper are based on a subset of the full 11 year data release, which includes the 34 pulsars with a timing baseline greater than three years. This restriction is justifiable since we do not expect any detectable GW signal to be present at frequencies $\gtrsim 3\ {\rm yr}^{-1}$, and it has the advantage of making our spatially correlated analysis—required to search for H.–D. correlations in the residuals—more computationally tractable, since the computational cost scales roughly as the cube of the number of pulsars. Table 3 lists the 34 pulsars with their epoch-averaged rms residuals, number of epochs and TOAs, and timing baselines.

As discussed in Section 3.4, we perform analyses for variants of our data model that reflect different assumptions about common red-spectrum processes, as listed in Table 1, and under four JPL ephemerides as well as BayesEphem (in select cases we include also the French INPOP13, which yields results broadly similar to DE430).

[57] https://github.com/stevertaylor/NX01
[58] https://github.com/jellis18/PAL2
[59] https://github.com/nanograv/enterprise
[60] https://github.com/jellis18/PTMCMCSampler
[61] https://github.com/nanograv/11yr_stochastic_analysis

Table 2
Prior Distributions Used in All Analyses Performed in This Paper

Parameter	Description	Prior	Comments
		White Noise	
E_k	EFAC per backend/receiver system	Uniform [0, 10]	Single-pulsar analysis only
Q_k [s]	EQUAD per backend/receiver system	Log-uniform [−8.5, − 5]	Single-pulsar analysis only
J_k [s]	ECORR per backend/receiver system	Log-uniform [−8.5, − 5]	Single-pulsar analysis only
		Red Noise	
$A_{\rm red}$	Red-noise power-law amplitude	Uniform [10^{-20}, 10^{-11}] (upper limits)	
		log-uniform [−20, − 11] (model comparison)	One parameter per pulsar
$\gamma_{\rm red}$	Red-noise power-law spectral index	Uniform [0, 7]	One parameter per pulsar
		BayesEphem	
$z_{\rm drift}$ [rad yr^{-1}]	Drift rate of Earth's orbit about the ecliptic z-axis	Uniform [-10^{-9}, 10^{-9}]	One parameter for PTA
$\Delta M_{\rm Jupiter}$ [$M_\odot$]	Perturbation to Jupiter's mass	$\mathcal{N}(0, 1.55 \times 10^{-11})$	One parameter for PTA
$\Delta M_{\rm Saturn}$ [$M_\odot$]	Perturbation to Saturn's mass	$\mathcal{N}(0, 8.17 \times 10^{-12})$	One parameter for PTA
$\Delta M_{\rm Uranus}$ [$M_\odot$]	Perturbation to Uranus' mass	$\mathcal{N}(0, 5.72 \times 10^{-11})$	One parameter for PTA
$\Delta M_{\rm Neptune}$ [$M_\odot$]	Perturbation to Neptune's mass	$\mathcal{N}(0, 7.96 \times 10^{-11})$	One parameter for PTA
PCA$_i$	ith PCA component of Jupiter's orbit	Uniform [−0.05, 0.05]	Six parameters for PTA
		Monopole-correlated Clock-error Signal, Power-law Spectrum	
$A_{\rm mono}$	Equivalent strain amplitude	Uniform [10^{-18}, 10^{-11}] (upper limits)	
		Log-uniform [−18, −14] (model comp., $\gamma = 13/3$)	One parameter for PTA
		Log-uniform [−18, −11] (model comp., γ varied)	One parameter for PTA
$\gamma_{\rm mono}$	GWB power-law spectral index	Delta function	Fixed, depends on analysis
		Dipole-correlated SSE error Signal, Power-law Spectrum	
$A_{\rm dip}$	Equivalent strain amplitude	Uniform [10^{-18}, 10^{-11}] (upper limits)	
		Log-uniform [−18, −14] (model comp., $\gamma = 13/3$)	One parameter for PTA
		Log-uniform [−18, −11] (model comp., γ varied)	One parameter for PTA
$\gamma_{\rm dip}$	GWB power-law spectral index	Delta function	Fixed, depends on analysis
		GWB, Power-law Spectrum	
$A_{\rm GWB}$	GWB strain amplitude	Uniform [10^{-18}, 10^{-11}] (upper limits)	
		Log-uniform [−18, −14] (model comp., $\gamma_{\rm GWB} = 13/3$)	One parameter for PTA
		Log-uniform [−18, −11] (model comp., $\gamma_{\rm GWB}$ varied)	One parameter for PTA
$\gamma_{\rm GWB}$	GWB power-law spectral index	Delta function	Fixed, depends on analysis
		GWB, Free Spectrum	
ρ_i [s^2]	GWB power-spectrum coefficients at $f = i/T$	Uniform in $\rho_i^{1/2}$ [10^{-18}, 10^{-8}][a]	One parameter per frequency
		GWB, Broken Power-law Spectrum	
$A_{\rm GWB}$	GWB broken power-law amplitude	Log-normal	One parameter for PTA
		$\mathcal{N}(-14.4, 0.26)$	MOP14
		$\mathcal{N}(-15, 0.22)$	S13
		$\mathcal{N}(-14.95, 0.12)$	Simon & Burke-Spolaor (2016) [b]
		$\mathcal{N}(-14.82, 0.08)$	Simon & Burke-Spolaor (2016) [c]
$\gamma_{\rm GWB}$	GWB power-law spectral index	Delta function	fixed to 13/3
κ	GWB broken power-law low-freq. spectral index	Uniform [0,7]	One parameter for PTA
$f_{\rm bend}$ [Hz]	GWB broken power-law bend frequency	Log-uniform [−9,−7]	One parameter for PTA
		GWB, Gaussian-process-interpolated Spectrum	
ρ_i [s^2]	GWB power-spectrum coefficients at $f = i/T$	$\mathcal{N}(0, V(\alpha_{\rm BH}, \rho_{\rm stars}, e_0))$	One parameter per frequency
$\alpha_{\rm BH}$	y-intercept of $M_{\rm BH}$–$M_{\rm bulge}$ relation	Uniform [7, 9]	One parameter for PTA
$\rho_{\rm stars}$ [$M_\odot$pc^{-3}]	Mass density of galactic-core stars	Log-uniform [1, 4]	One parameter for PTA
e_0	Binary eccentricity at formation	Uniform [0, 0.95]	One parameter for PTA

Notes.
[a] The uniform $\rho_i^{1/2}$ prior is chosen to be consistent with the uniform $A_{\rm GWB}$ prior for the power-law model, since $\varphi_{ii} \propto A_{\rm GWB}^2$.
[b] Uses the McConnell & Ma (2013) $M_{\rm BH}$–$M_{\rm bulge}$ relation.
[c] Uses the Kormendy & Ho (2013) $M_{\rm BH}$–$M_{\rm bulge}$ relation.

Table 3
Pulsars Used in Our GWB Analysis (see NG11 for Full Details of Pulsar Properties)

PSR Name	rms (μs)	#Epochs	#TOAs	Baseline (yr)
J0023+0923	0.361	415	8217	4.4
J0030+0451	0.691	268	5699	10.9
J0340+4130	0.454	127	6475	3.8
J0613−0200	0.422	324	11566	10.8
J0645+5158	0.178	166	6370	4.5
J1012+5307	1.07	493	16782	11.4
J1024−0719	0.323	194	8233	6.2
J1455−3330	0.672	277	7526	11.4
J1600−3053	0.23	275	12433	8.1
J1614−2230	0.199	241	11173	7.2
J1640+2224	0.426	323	5982	11.1
J1643−1224	3.31	298	11528	11.2
J1713+0747	0.108	789	27571	10.9
J1738+0333	0.52	208	4881	6.1
J1741+1351	0.128	134	3047	6.4
J1744−1134	0.842	322	11550	11.4
J1747−4036	3.59	113	6065	3.8
J1853+1303	0.239	107	2514	4.5
B1855+09	0.809	296	5634	11.0
J1903+0327	3.65	112	3326	6.1
J1909−3744	0.148	451	17373	11.2
J1910+1256	0.544	130	3563	6.8
J1918−0642	0.322	364	12505	11.2
J1923+2515	0.229	87	1954	4.3
B1937+21	1.57	460	14217	11.3
J1944+0907	0.352	104	2850	4.4
B1953+29	0.377	88	2331	4.4
J2010−1323	0.257	222	10844	6.2
J2017+0603	0.11	102	2359	3.8
J2043+1711	0.12	197	3262	4.5
J2145−0750	0.968	258	10938	11.3
J2214+3000	1.33	176	4569	4.2
J2302+4442	1.07	138	6549	3.8
J2317+1439	0.271	395	5958	11.0

Note. The second column shows the weighted root mean square epoch-averaged post-fit timing residuals (see NG9a for a definition of this quantity).

4.1. Bayesian Upper Limits

Following NG9b, we present upper limits on the strain amplitude of a GWB modeled as a power law and as a free spectrum (see Section 3.2).

We first discuss our limit on the GWB from a population of GW-driven SMBHB inspirals, as represented by Equation (5) for $\gamma = 13/3$. Adopting each of the JPL ephemerides as fixed-parameter models and including H.–D. interpulsar correlations in the likelihood (model 3A), the 95% upper limit on $A_{\rm GWB}$ ranges from $1.53(1) \times 10^{-15}$ for DE421 to $1.78(1) \times 10^{-15}$ for DE430, where the digits in parentheses give 1σ uncertainties (see the third bolded column of Table 4). Indeed, the choice of ephemeris leads to systematic biases that are larger than the statistical uncertainty of the limits. Furthermore, the limits do not evolve monotonically with later ephemerides, suggesting that even DE436, the most recent ephemeris based on the most data, is still measurably displaced from truth for the purposes of our analysis.

We therefore chose to marginalize the $A_{\rm GWB}$ posterior over ephemeris uncertainties, using the BayesEphem model described in Section 3.3. Doing so yields our fiducial 11 year upper limit of $1.45(2) \times 10^{-15}$. This value is the same, within sampling error, no matter which ephemeris we take as a starting point for BayesEphem, demonstrating that we have successfully "bridged" the individual ephemerides.

Comparing the columns of Table 4 shows how the upper limits vary under different assumptions on the presence of spatially correlated common processes in the data. The limits are slightly more stringent if we model the GWB as a spatially uncorrelated common process (model 2A in the second column), indicating that H.–D. correlations help the likelihood isolate a GW-like signal (whether real or due to random noise fluctuations). Introducing additional spatially correlated processes (with ephemeris-error-like dipolar correlations, clock-error-like monopolar correlations, or both, corresponding to models 3B, 3D, and 3C) reduces upper limits for the individual ephemerides but not for BayesEphem, suggesting that the same realization of interpulsar signal correlations can be picked up by different ORFs, and that dipole and monopole processes can absorb some, but not all, of the systematic bias caused by ephemeris error.

In Figure 2 we show the 95% upper limit for the amplitude of an uncorrelated common process (model 2A) as a function of γ. In the absence of red noise, and if the lowest sampling frequency ($1/T$) dominated our sensitivity, we would expect these constraints to scale as $\propto T^{-\gamma/2}$, where T is the longest timing baseline across the entire PTA. We find the actual scaling to be closer to $\propto T^{-0.4\gamma}$, indicating that red noise is present and that more than one frequency component contributes to the likelihood.

In the top panel of Figure 3, we show 95% upper limits for free-spectrum amplitudes (jagged black line), which are diagnostics of the sensitivity of our data set to individual monochromatic GW signals. In the same plot we show also the $\gamma = 13/3$ ($\alpha = -2/3$) power-law limit (straight black line). The thickness of the lines indicates the spread of results over ephemeris choices. Sensitivity is inhibited at lower frequencies by fitting pulsar spin-down parameters and is dominated at higher frequencies by white noise, matching the expected $f^{3/2}$ expected slope shown as the dashed–dotted line. Sensitivity is also noticeably lost at $f = {\rm yr}^{-1}$ by fitting pulsar positions. The colored dashed lines and bands display a representative selection of theoretical expectations for the SMBHB GWB: McWilliams et al. (2014, hereafter MOP14), Simon & Burke-Spolaor (2016), and Sesana et al. (2016, hereafter S16). The models in Simon & Burke-Spolaor (2016) and S16 are broadly similar, differing predominantly in the choice of SMBH–host-galaxy mass relationship, where S16 factors in potential selection biases in dynamically measured SMBH masses (Shankar et al. 2016). These same results and theoretical expectations are shown in the bottom panel of Figure 3 in terms of the stochastic GWB energy density (per logarithmic frequency bin) in the universe as a fraction of closure density, $\Omega_{\rm GWB}(f)h^2$, where the scaling by h^2 makes the results agnostic of the specific value of the Hubble constant. The fractional energy density scales as $\Omega_{\rm GWB}h^2 \propto f^2 h_c(f)^2$. In Section 5.1 below we present an astrophysical discussion of our results.

4.2. Bayesian Model-comparison Evidence for GWs

In Tables 5 and 6 and in Figure 4, we show Bayes factors for two sets of model comparisons performed on the 11 year data set to quantify the statistical evidence for a stochastic GWB and for coherent sources of systematic errors that lead to spatially correlated residuals. The first four columns of Table 5 and the

Table 4

GWB Amplitude 95% Upper Limits for the NANOGrav 11 year Data Set, Computed for a Power-law Spectrum with $\gamma = 13/3$, and with Uniform Prior on $A_{\rm GWB}$ (see Equations (4) and (5))

JPL Ephemeris	95% Upper Limit on $A_{\rm GWB}$ [$\times 10^{-15}$], $\gamma = 13/3$ Power Law				
	Uncorrelated Common Process (2A)	H.–D. Correlated Common Process			
		Alone (3A)	+ Dipole (3B)	+ Monopole, Dipole (3C)	+ Monopole (3D)
DE421	1.505(8)	**1. 53(1)**	1.478(8)	1.487(8)	1.53(3)
DE430	1.76(2)	**1. 79(1)**	1.698(9)	1.676(9)	1.74(2)
DE435	1.57(3)	**1. 60(1)**	1.555(8)	1.55(1)	1.58(2)
DE436	1.61(2)	**1. 67(1)**	1.594(9)	1.56(1)	1.60(2)
INPOP13c	1.74(3)	…	…	…	…
BayesEphem	1.34(1)	***1.45(2)***	1.52(3)	1.49(3)	1.48(4)

Note. We report limits for an uncorrelated common process (as in NG9b) and for a Hellings–Downs spatially correlated process, either alone (in bold, our fiducial result) or in the presence of additional correlated processes with different ORF. "…" indicates that analyses for these numbers were not performed.

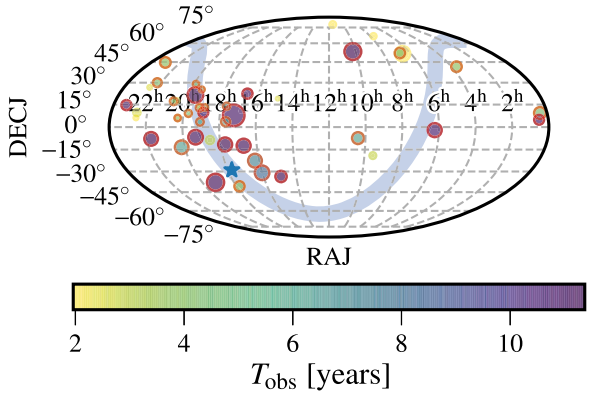


Figure 1. Sky positions of all 45 pulsars in the NANOGrav 11 year data set. The area of each circle is indicative of the number of TOAs, while the color scale indicates the observational baseline. The 34 pulsars whose baselines are longer than three years are indicated with solid red edges. The Milky Way plane is shown behind as a blue band (thickness is not indicative of Galactic scale height), with the Galactic center shown as a blue star. The longest baseline is given by J1744–1134 with 11.37 years, while the largest data set is given by J1713+0747 with 27,571 TOAs.

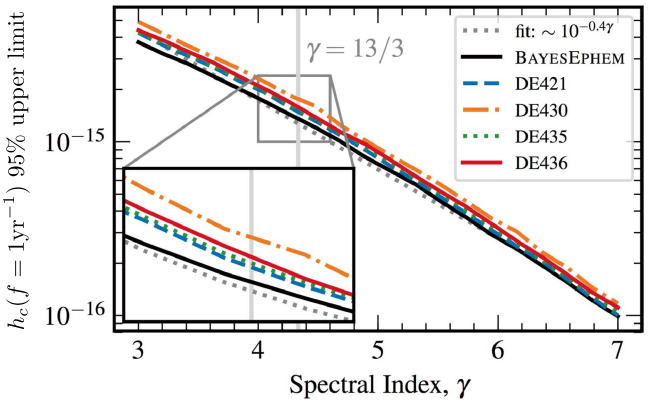


Figure 2. GWB amplitude 95% upper limit for an uncorrelated common process (model 2A) as a function of spectral index γ (see Equation (5)) for the JPL ephemerides and for BayesEphem. The dotted curve shows a power-law fit to the BayesEphem curve, which is consistent with a similar fit in NG9b.

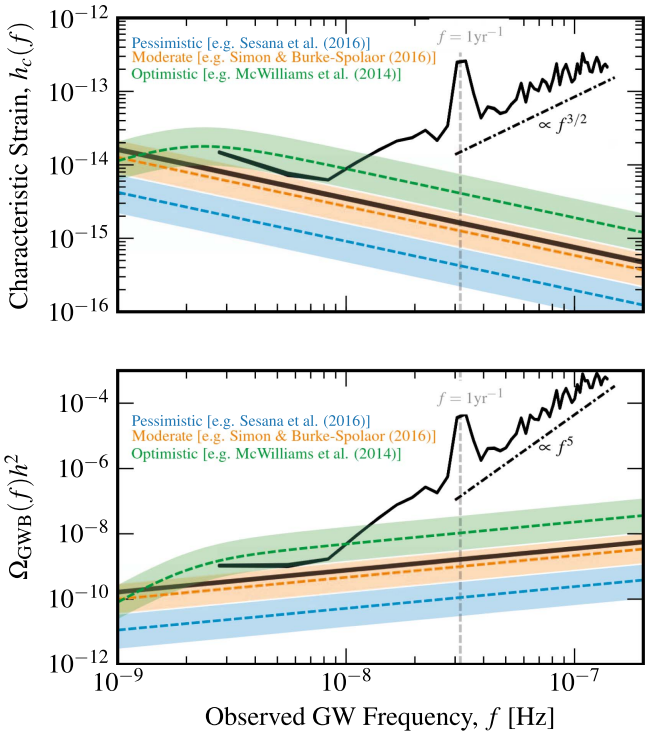


Figure 3. Top panel: GWB amplitude 95% upper limits for an uncorrelated common process with a $\gamma = 13/3$ power law (straight black line) or with independently determined free-spectrum components (jagged black line). The thickness of the lines spans the spread of results over different ephemerides. The dashed–dotted line shows the expected sensitivity scaling behavior for white noise. The colored dashed lines and bands show the median and one-sigma ranges for the GWB amplitudes predicted in MOP14 (green), Simon & Burke-Spolaor (2016; orange), and S16 (blue). Bottom panel: as in the top panel, except showing the results in terms of the stochastic GWB energy density (per logarithmic frequency bin) in the universe as a fraction of closure density, $\Omega_{\rm GWB}(f)h^2$. The relationship between $h_c(f)$ and $\Omega_{\rm GWB}(f)h^2$ is given in Equation (10).

graph on the left of Figure 4 are diagnostics of the multilevel decision scheme outlined above in Section 3.1. Adopting the JPL ephemerides as fixed-parameter models, the data favor the presence of a common uncorrelated process in all pulsars to various degrees and especially so for DE430, and they favor slightly the presence of H.–D. interpulsar correlations. However, this preference disappears if we marginalize over the ephemeris uncertainties.

The effects of ephemeris errors are also apparent in the upper plot of Figure 5, which shows the posterior distribution of $\log_{10}A_{\rm GWB}$ under the log-uniform prior used to compute Bayes factors, for $\gamma = 13/3$, and ignoring H.–D. correlations. The dashed lines show the posterior obtained by taking each ephemeris as fixed-parameter models without uncertainties; the

Table 5
Bayes Factors for Model Comparisons Using NANOGrav's 11 year Data Set, Performed to Examine the Evidence for a GWB

JPL Ephemeris	Uncorr. Red Process Versus Pulsar Noise		H.–D. Corr. Red Process Versus Pulsar Noise							
	(2A–1)		(3A–1)		With Dipole (3B–2B)		With Dipole, Monopole (3C–2C)		With Monopole (3D–2D)	
	$\gamma = 13/3$	$\gamma \in [0, 7]$	$\gamma = 13/3$	$\gamma \in [0, 7]$	$\gamma = 13/3$	$\gamma \in [0, 7]$	$\gamma = 13/3$	$\gamma \in [0, 7]$	$\gamma = 13/3$	$\gamma \in [0, 7]$
DE421	8.28(4)	5.3(2)	11.9(7)	6.5(2)	3.57(5)	2.07(6)	3.20(5)	1.96(5)	7.4(5)	3.7(3)
DE430	18.9(7)	8.7(4)	26(2)	12.8(9)	3.69(4)	2.05(3)	3.94(9)	1.9(1)	12(1)	5.6(4)
DE435	1.82(4)	1.22(1)	2.15(4)	1.69(5)	1.52(2)	1.17(2)	1.35(2)	0.99(2)	1.77(4)	1.43(4)
DE436	4.4(1)	3.5(2)	7.2(4)	4.8(2)	2.17(4)	1.54(3)	2.14(2)	1.34(4)	3.4(1)	2.18(5)
INPOP13c	24.9(7)	…	…	…	…	…	…	…	…	…
BayesEphem	0.884(9)	0.647(7)	1.00(2)	0.70(1)	0.93(2)	0.67(3)	0.98(4)	0.66(2)	0.98(5)	0.70(3)

Note. The digit in parentheses gives the uncertainty on the last quoted digit. All factors were computed with the Savage–Dickey formula (Equation (2)), with the hyperparameter priors listed in Table 2.

solid lines show the posteriors obtained by marginalizing over ephemeris uncertainties, starting with each ephemeris in turn. Although the dashed curves agree roughly in their modes, they have surprisingly different widths and contrast, which we may define as the ratio of the peak posterior to its value in the lowest-amplitude (leftmost) bin; the latter appears in the denominator of the Savage–Dickey estimate (Equation (2)) of the signal-versus-noise Bayes factor.

The convergence of the solid lines to a flatter common shape demonstrates that our modeling of ephemeris uncertainties bridges the four ephemerides successfully, removing spurious evidence for GWs, or potentially absorbing a true GW signal. However, if a true GW signal is present, it happens to be significantly covariant with the systematic differences in the Roemer delays induced by the last few ephemerides; furthermore, the signal appears to weaken as we shift from older (DE421, DE430) to newer, plausibly more accurate ephemerides (DE435, DE436), although this trend is not entirely consistent. In this paper, we do not attempt to quantify whether these circumstances are realized often in the ensemble of possible data sets similar to ours; nevertheless, these circumstances motivate our choice of marginalizing over ephemeris uncertainties as the principled Bayesian strategy for our analysis.

The six rightmost columns of Table 5, as well as Table 6 and the graph on the right of Figure 4, document the degree to which the data favor the presence of timing-residual components with different spatial correlations. Components with both dipolar (ephemeris-error-like) and monopolar (clock-error-like) correlations are disfavored, although this conclusion is significantly weakened if we marginalize over ephemeris uncertainties. At the same time, the evidence for quadrupolar (GWB-like) correlations is weakened when the model allows for other spatially correlated processes. This is not unexpected, since spatial correlations with different multipolar structures only become truly orthogonal in the limit of many equally low-noise pulsars.

Indeed, discrimination of monopolar, dipolar, and quadrupolar correlation signatures will improve as our data sets gain more and more pairs of high-timing-precision pulsars with a broad distribution of angular separations. We plan to characterize discrimination requirements (on pulsar number, timing quality, and sky position) in our upcoming paper on SSE error modeling.

Impact of SSE Error Modeling on GW Detection—We performed a small number of simulations to test the impact of BayesEphem on our GWB detection prospects over the next few years. To this end, we produced realistic 15 year data sets[62] using DE436 and injecting GWBs of various amplitudes, and we analyzed the full data sets, as well as their 11.4 year "slices," using DE430 and BayesEphem. We chose DE430 because it led to the highest signal-versus-noise Bayes factor (model 2A-versus-1) and upper limits for the actual data.

For a noise-only simulation, we find that unmodeled systematic offsets between DE436 and DE430 are interpreted as a common red-spectrum process with a signal-versus-noise Bayes factor (model 2A-versus-1) of $\sim$2 in 11.4 years of data and $\sim$20 in 15 years of data. By contrast, BayesEphem is able to account for the offsets, reducing Bayes factors to levels consistent with noise fluctuations. As we increase the injected GWB amplitude, model 2A-versus-1 Bayes factors remain low for 11.4 years of data, even for amplitudes comparable to our fiducial upper limits. The same is true for model 3A-versus-2A Bayes factors (the definitive spatial-correlation test for GWBs), which are plotted in Figure 6.

For 15 years of data, the scaling of Bayes factors with injected GWB amplitude is comparable, for both DE430 and BayesEphem. Remarkably, the potential covariance of BayesEphem parameters with GWB amplitude does not inhibit signal detection in the near future, even at astrophysically pessimistic levels ($\sim 5 \times 10^{-16}$, consistent with Sesana et al. 2016). Thus, while SSE errors may spuriously produce early signs of a GWB (i.e., a common red-spectrum process), their mitigation with BayesEphem will not impair prospects for near-future GWB detection. We regard our simulations as conservative, since additional pulsars, as well as improved timing precision and SSE accuracy, will accelerate progress toward detection.

[62] To produce the data sets, we used actual observation epochs for the 34 NANOGrav pulsars and set residuals equal to white measurement noise plus red-spectrum intrinsic noise, at levels consistent with those estimated for the actual data (NG11). We rescaled TOA uncertainties by a factor of 1.5, which calibrates the noise-only simulated data set so that its 11.4 year "slice" has the same (DE436, model 2A) GWB upper limit as the real data. We extended the data set baseline to 15 years by drawing observation epochs and TOA measurement errors from distributions of these quantities over the last three years of real data.

Table 6
Bayes Factors for Model Comparisons Using NANOGrav's 11 Year Data Set, Performed to Examine the Evidence for Spatial Correlations with Different ORF

JPL Ephemeris	3A–2A: H.-D. Corr. Red Process		2B–2A: Dipole Corr. Red Process Versus Uncorrelated Red Process		2D–2A: Monopole Corr. Red Process	
	$\gamma = 13/3$	$\gamma \in [0,7]$	$\gamma = 13/3$	$\gamma \in [0,7]$	$\gamma = 13/3$	$\gamma \in [0,7]$
DE421	1.34(7)	1.53(8)	0.46(3)	0.60(3)	0.18(1)	0.21(1)
DE430	1.44(8)	1.7(1)	0.46(3)	0.94(6)	0.106(9)	0.21(2)
DE435	1.24(6)	1.42(7)	0.55(3)	0.85(4)	0.54(3)	0.55(3)
DE436	1.45(8)	1.63(9)	0.57(3)	1.05(6)	0.46(3)	0.52(3)
BayesEphem	1.08(6)	1.15(7)	0.83(5)	0.87(6)	1.12(9)	0.96(7)

Note. The digit in parentheses gives the uncertainty on the last quoted digit. All factors were computed with the product-space method discussed in Section 3.1, with the hyperparameter priors listed in Table 2.

4.3. Optimal Statistic

Table 7 compares the fixed-noise and noise-marginalized optimal statistic (see Section 3.5) for an H.–D. spatially correlated common process computed using DE421, DE430, DE435, and DE436. The noise marginalization was performed using 10,000 realizations of the noise. Except for DE421, the fixed-noise analysis systematically underestimates $\hat{A}^2_{\rm GWB}$ and S/N compared to the noise-marginalized analysis because of the covariance between pulsar red-noise parameters and the common red-noise parameters. Note that, although the optimal statistic is formulated in terms of the squared amplitude, negative $\hat{A}^2_{\rm GWB}$ and S/N values are possible if noise fluctuations result in negative correlations. In the noise-marginalized analysis, we find mean S/N <1 for all ephemerides—there is no appreciable evidence of H.–D. correlations. These results are consistent with the Bayesian analysis.

Table 8 compares the noise-marginalized optimal statistic computed for H.–D. spatial correlations with variants of the statistic that model dipolar and monopolar correlations. In addition to computing the optimal statistic using individual ephemerides, we also use BayesEphem to marginalize over the ephemeris uncertainty. For all of these analyses, we find no evidence for a common process with either H.–D., monopolar, or dipolar spatial correlations.

The upper half of Figure 7 shows the mean noise-marginalized cross-correlated power between pulsar pairs as a function of angular distribution, averaged into 10 degree bins. There is no evidence of the H.–D. correlations characteristic of isotropic GWBs. The lower half of the plot shows a histogram of angular separations for the pulsar pairs in our data set: NANOGrav is currently most sensitive to angular separations between 30° and 60°, which correspond to the smallest errors in the cross-correlation plot.

4.4. Comparison of 9 Year and 11 Year Results

The nine-year analysis of NG9b adopted DE421 as a fixed-parameter model without uncertainties and did not include H.–D. correlations. Thus, a straight comparison can be made with the 11 year DE421 model-2A results: the $\gamma = 13/3$ upper limit remains at 1.5×10^{-15}, while the $\gamma = 13/3$ Bayes factor versus pulsar noise changes from 0.81 to 8.3; however, this comparison is not very significant given what we have learned about ephemeris errors.

Applying BayesEphem to the nine-year data set successfully bridges $A_{\rm GWB}$ posteriors (see Figure 8) and yields a model-2A (uncorrelated) upper limit of $2.67(2) \times 10^{-15}$ and a model-3A (H.–D.-correlated) upper limit of $2.91(2) \times 10^{-15}$ (both for $\gamma = 13/3$). Thus, our fiducial model-3A upper limit improves by a factor of $2.91/1.45 = 2.0$ in the 11 year data set. This is greater than expected from simple scaling arguments (Siemens et al. 2013), for which the additional two years of data should reduce the limit from 2.91×10^{-15} to 1.85×10^{-15}, i.e., an improvement of $\sim$1.6. The major cause of this discrepancy is presumably the longer 11 year baseline being better able to disentangle ephemeris perturbations, which have typical timescales of the 11.86 year Jupiter period.

The model-2A Bayes factors versus pulsar noise are 0.910(7) for $\gamma = 13/3$ and 1.210(4) for $\gamma \in [0, 7]$, while they are 1.27 (1) and 2.29(3) for model 3A. All Bayes factors under BayesEphem are comparably uninformative for the nine-year and 11 year data sets.

We also reproduce the spectral-turnover analysis of NG9b, which models the GWB with a broken power-law spectrum (Equation (6), following Sampson et al. 2015). Figure 9, obtained with $A_{\rm GWB}$ priors from MOP14, Simon & Burke-Spolaor (2016), and S16 (using BayesEphem and ignoring H.–D. correlations) can be contrasted with Figure 5 of NG9b. In both figures, models that infer larger GWB levels require turnovers at higher GW frequencies to be consistent with the data: the MOP14 prior gives a median value of $f_{\rm bend}$ at 3.83×10^{-8} Hz, while the S16 prior gives a value of 1.09×10^{-8} Hz, a difference of more than a factor of three. This analysis is a useful tool in broadly understanding various models' consistency with PTA limits. However, it is limited by attributing a single value of κ (see Equation (6)) to the entire population of SMBHBs, and it is unable to incorporate eccentricity, which flattens out the turnover and skews toward higher GW frequencies. In Section 5, we present results incorporating a more sophisticated approach (Taylor et al. 2017b), allowing us to confront astrophysical population models directly.

5. Limits on Astrophysical Models

Some of the most exciting science made possible by the NANOGrav data is realized when we use the GWB constraints to confront the astrophysics of various source populations. Now, the most likely source population for PTAs is SMBHBs. In NG9b, we introduced simple PTA constraints on SMBHB population parameters, but due to methodological limitations

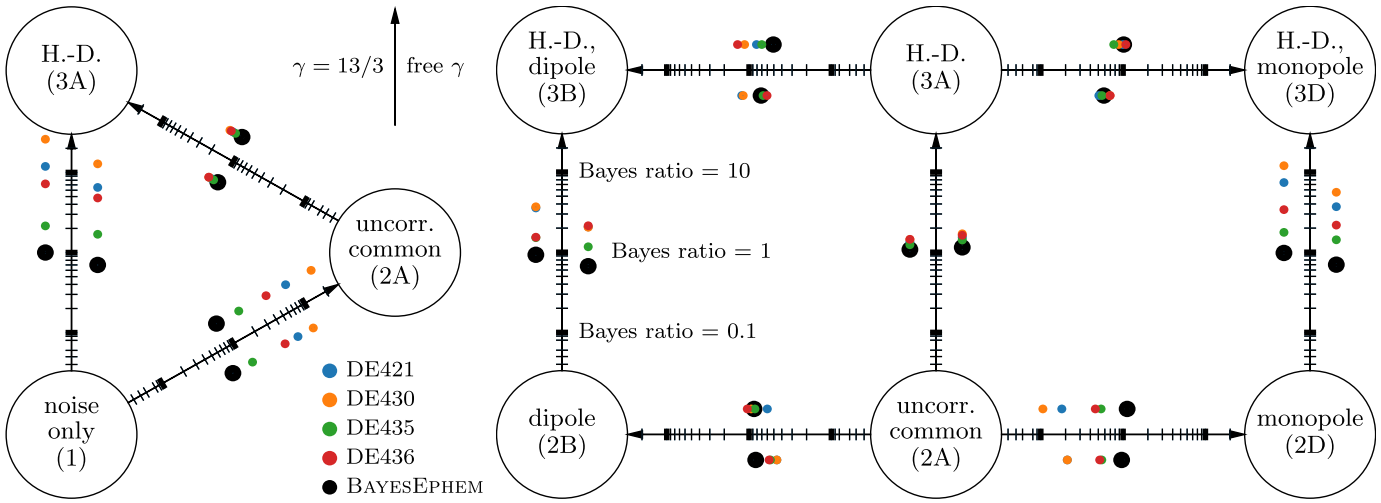


Figure 4. Bayes factors for model comparisons on the 11 year data set: on the left, evidence of a GWB, on the right, effects of spatially correlated systematics. In these graphs, each model (as described in Table 1) is represented by a bubble, and for each pair of models the dots mark on a logarithmic scale the measured Bayes factor in favor of the model at the head of the arrow. Thus, dots are closer to the model favored by the data. The smaller colored dots represent Bayes factors computed by taking one of the DE421, DE430, DE435, and DE436 JPL ephemerides as a fixed-parameter model without uncertainties; the larger black dots represent Bayes factors computed by marginalizing over ephemeris errors (i.e., by adopting BayesEphem). Dots to the left of the arrows correspond to fixing the spectral slope γ of the GWB to 13/3, as appropriate for a background from SMBHBs evolving purely by GW emission; dots to the right correspond to marginalizing over γ, taken to have uniform prior distribution in [0, 7]. The graph on the left shows that when adopting the JPL ephemerides as fixed-parameter models, most of the evidence for a GWB accrues from the presence of unexplained red-spectrum residuals in each pulsar (2A–1), with a smaller preference added by modeling Hellings–Downs correlations (3A–2A); neither conclusion is supported by BayesEphem. As for the graph on the right, the bottom row compares a common uncorrelated red process with dipolar and monopolar processes; the former is favored. The top row examines the case for dipolar and monopolar processes in the presence of a Hellings–Downs (GW-like) signal. Comparing the vertical arrows in the left and right graphs, we see that (for fixed JPL ephemerides) evidence for a GW-like signal is weakened when the model allows for other spatially correlated processes.

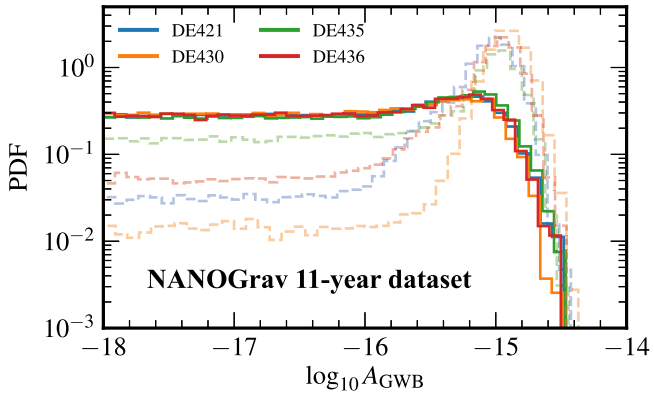


Figure 5. Posterior probability distributions for $A_{\rm GWB}$ (log-uniform prior, $\gamma = 13/3$, and no spatial correlations), as computed for the NANOGrav 11 year data set under individual JPL ephemerides (dashed lines), and with BayesEphem, taking each of the JPL ephemerides as a starting point (solid lines). This plot demonstrates that BayesEphem bridges the JPL ephemerides successfully; in doing so it removes most evidence for the presence of a GWB.

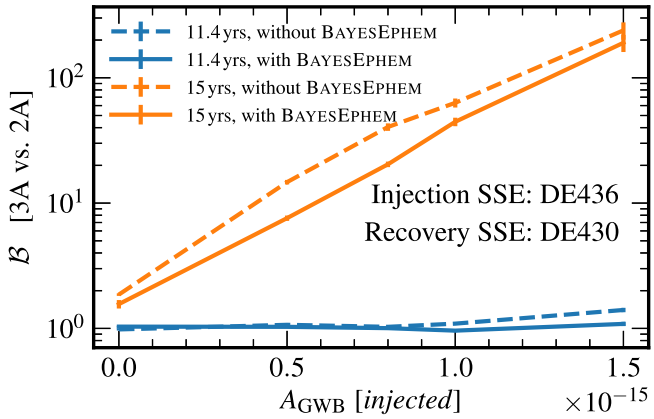


Figure 6. Model 3A-versus-2A (spatial correlation) Bayes factors for a set of simulated 15 year data sets with GWB injections at different levels. We analyze the full data sets (orange lines) as well as their 11.4 year "slices" (blue), both by adopting the "wrong" ephemeris (dashed) and by employing BayesEphem to marginalize over SSE errors (solid). We conclude that BayesEphem will not impede the ability of PTAs to make a definitive detection in the near future (see main text).

we were unable to deliver a realistic analysis, i.e., we derived constraints for the parameters describing a broken power-law spectrum and then reinterpreted those constraints in terms of SMBHB effects that could alter the spectrum, taken one at a time. In this paper, we adopt the modeling framework developed by Taylor et al. (2017b) to go much further: we use a set of population-synthesis simulations to explore the effects of population parameters on the GWB spectrum then constrain those population parameters directly from the data. We also apply the same method to the most recent cosmic-string models.

5.1. Supermassive Black Hole Binaries

PTAs are sensitive to the stochastic GWB comprised of the superposition of GWs from merging SMBHBs throughout the universe. The details of this background (i.e., spectral shape and amplitude) are sensitive to the physics of SMBHB evolution. The history of SMBHB mergers is generally assumed to follow the history of galaxy mergers, but the exact relation remains an open question. Dynamical friction initially causes SMBHs to sink toward each other in a postmerger galaxy remnant, but becomes an inefficient means of further hardening at parsec separations (Begelman et al. 1980).

Table 7
Optimal Statistic $\hat{A}^2_{\rm gw}$ and Associated S/N for the NANOGrav 11 Year Data Set, Assuming a $\gamma = 13/3$ Power-law GWB with Hellings–Downs Spatial Correlations

	Fixed Noise		Noise Marginalized	
JPL Ephemeris	$\hat{A}^2_{\rm GWB}$	S/N	Mean $\hat{A}^2_{\rm GWB}$	Mean S/N
DE421	8.23×10^{-31}	1.06	8.9×10^{-31}	0.9(9)
DE430	2.32×10^{-31}	0.390	6.9×10^{-31}	0.4(4)
DE435	-3.46×10^{-31}	−0.640	5.9×10^{-31}	0.7(6)
DE436	4.47×10^{-32}	−0.069	9.7×10^{-31}	0.8(7)

Note. The noise-marginalized computation provides a more accurate assessment of the significance of a common red process compared to the fixed noise due to the covariance between pulsar red-noise parameters and common red-noise parameters (see Section 3.5).

Additional dynamical influences are required to drive an SMBHB to milliparsec orbital separations, and thus into the PTA frequency band. This supposed final parsec problem can be overcome by a variety of processes: (i) three-body scattering effects with stars in the galaxy's bulge, where stars in the binary's loss cone slingshot off the binary, carrying away orbital energy (Mikkola & Valtonen 1992; Quinlan 1996; Sesana et al. 2006), (ii) interactions between the binary and a viscous circumbinary disk (Ivanov et al. 1999; Haiman et al. 2009; Kocsis & Sesana 2011), and (iii) eccentricity, which increases the rate of binary evolution (Peters & Mathews 1963; Peters 1964) and can be amplified by (i) (Sesana et al. 2006; Sesana 2010; Roedig & Sesana 2012) and (ii) (Armitage & Natarajan 2005; Cuadra et al. 2009; Roedig et al. 2011). If the final parsec problem is not completely overcome by additional environmental processes, a subsequent galaxy merger may add a third massive black hole to the system, which can drive the initial binary toward coalescence and may increase that binary's eccentricity (Bonetti et al. 2017; Ryu et al. 2018). All of these influences can cause the shape of the GWB spectrum in the PTA band to deviate from the fiducial $f^{-2/3}$ power law at low frequencies ($f \in [1, 10]$ nHz), causing a change in slope or a turnover if the binary remains coupled to the environment or has large orbital eccentricities (Enoki et al. 2007; Kocsis & Sesana 2011; Sesana 2013; Ravi et al. 2014; Huerta et al. 2015; Sampson et al. 2015; Taylor et al. 2017b).

As discussed earlier, we use a GP spectral model to explore the parameter space of SMBHB environments and dynamics. We perform sophisticated population-synthesis simulations over a $5 \times 5 \times 5$ grid in the $\{\alpha_{\rm BH}, \rho_{\rm stars}, e_0\}$ parameter space, where $\alpha_{\rm BH}$ is the y-intercept of the $M_{\rm BH}$–$M_{\rm bulge}$ relationship, $\rho_{\rm stars}$ is the typical mass density of galactic-core stars at the binary influence radius, and e_0 is binary eccentricity at formation. At each grid point, we perform 100 simulations and compute the mean spectrum and uncertainty from Poisson variation. We then train a GP at each GW frequency, allowing spectral amplitudes to be predicted with uncertainties over the entire parameter space. These predictions act as priors on the strain within the free-spectrum model. We set uniform priors on the astrophysical parameters corresponding to $\alpha_{\rm BH} \in \{7, 9\}$, $\log_{10}[\rho_{\rm stars}/M_\odot {\rm pc}^{-3}] \in \{1, 4\}$, and $e_0 \in \{0, 0.95\}$.

Our population-synthesis model is similar to the scheme described in Simon & Burke-Spolaor (2016) and Taylor et al. (2017b), where the SMBHB merger rate density was constructed from observed galaxy properties and SMBH–host-galaxy relations. Specifically, we adopt a galaxy stellar mass function from Ilbert et al. (2013), a galaxy pairing fraction from Robotham et al. (2014), and a parametrized $M_{\rm BH}$–$M_{\rm bulge}$ relationship. The $M_{\rm BH}$–$M_{\rm bulge}$ relation is set by three parameters: $\alpha_{\rm BH}$, $\beta_{\rm BH}$, and $\epsilon_{\rm BH}$, where $\log_{10}(M_{\rm BH}/M_\odot) = \alpha_{\rm BH} + \beta_{\rm BH} \log_{10}(M_{\rm bulge}/10^{11} M_\odot)$ and $\epsilon_{\rm BH}$ is the intrinsic scatter of points around the set power law. We fix $\beta_{\rm BH} = 1$ and $\epsilon_{\rm BH} = 0.3$, values typical of observational measurements (see, e.g., Kormendy & Ho 2013; McConnell & Ma 2013). As shown in Simon & Burke-Spolaor (2016), $\alpha_{\rm BH}$ is the parameter of maximal impact on the GWB; as such, it is the only parameter from the $M_{\rm BH}$–$M_{\rm bulge}$ relation that is varied in this work. However, there are impacts on the level of the GWB predicted from changing all of these parameters, which are explored in depth in Simon & Burke-Spolaor (2016), and therefore the limits on $\alpha_{\rm BH}$ in this work must be interpreted in that context.

The eccentricity evolution in this model follows the prescription first derived in Quinlan (1996) and later expanded upon in Sesana (2010). However, recent work in Rasskazov & Merritt (2017a; see also Sesana et al. 2011; Gualandris et al. 2012; Mirza et al. 2017) has shown that eccentricity evolution can be damped by the rotation of the central stellar bulge, which would lessen the effect of extreme initial eccentricities.

The parameter $\alpha_{\rm BH}$ primarily changes the overall level of the GWB, while $\rho_{\rm stars}$ and e_0 primarily change its spectral shape. We start to explore this parameter space by constraining $\alpha_{\rm BH}$. Figure 10 shows $\alpha_{\rm BH}$ posteriors derived by marginalizing over $\{\rho_{\rm stars}, {\rm e}_0\}$ (black solid line) and by marginalizing over $\rho_{\rm stars}$ for circular sources ($e_0 = 0$; red dashed line). The constraint for circular sources is slightly more stringent, as is expected from removing a degree of freedom. However, in both cases the determination of $\alpha_{\rm BH}$ in Kormendy & Ho (2013, hereafter KH13) is disfavored compared to the others.

Quantitatively, we may take the ratios of PDFs as proxies for model-comparison Bayes factors between $\alpha_{\rm BH}$ determinations; by doing so, we find McConnell & Ma (2013, hereafter MM13) to be 1.5 times more probable than KH13, while the other two measurements are 1.9 times more probable than KH13. These constraints become 2 and 2.6, respectively, for circular sources ($e_0 = 0$). As stated above, these results for $\alpha_{\rm BH}$ need to be viewed in the context of the complete model used to infer the population of SMBHBs, which relies on the assumption that the SMBH merger rate follows the observed galaxy merger rate. This may not be the case if the final parsec problem is not solved for all systems, or if binary evolution takes much longer than anticipated by this model (Tremmel et al. 2018). However, even when we robustly incorporate many of the parameters that impact the spectral shape of the GWB, the NANOGrav 11 year data set prefers values of $\alpha_{\rm BH}$ that are lower than the largest observed measurements from KH13.

We can also compute a joint marginalized posterior for $\rho_{\rm stars}$ and e_0, but the $\alpha_{\rm BH}$ distribution is too broad for this to be useful. It is more informative to examine $p(\rho_{\rm stars}, e_0|\alpha_{\rm BH})$ for a few representative values of $\alpha_{\rm BH}$. In Figure 11, we show $\alpha_{\rm BH} = 8.46$, 8.69, and 8.89. The first two values are the measurements reported in MM13 and KH13, while the third is an even larger value. The top panels of Figure 11 show posteriors, while the bottom panels show the corresponding marginalized spectral distributions using the same conventions as Figure 9. As $\alpha_{\rm BH}$ increases from left to right, the GWB increases in level, and its spectral shape needs to deviate more strongly from an $f^{-2/3}$ power law to be consistent with the

Table 8

Noise-marginalized Optimal Statistic $\hat{A}^2_{\rm gw}$ and Associated S/N for the NANOGrav 11 Year Data Set, Assuming a $\gamma = 13/3$ Power-law GWB with Hellings–Downs (GW-like), Monopolar (Clock-error-like), and Dipolar (Ephemeris-error-like) Spatial Correlations

	Hellings–Downs		Monopole		Dipole	
JPL Ephemeris	Mean $\hat{A}^2_{\rm GWB}$	Mean S/N	Mean $\hat{A}^2_{\rm GWB}$	Mean S/N	Mean $\hat{A}^2_{\rm GWB}$	Mean S/N
DE421	8.9×10^{-31}	0.9(9)	-6.2×10^{-33}	0.0(6)	3.8×10^{-32}	0(1)
DE430	6.9×10^{-31}	0.4(4)	-1.5×10^{-31}	0.5(4)	2.4×10^{-31}	0.7(9)
DE435	5.9×10^{-31}	0.7(6)	8.5×10^{-32}	0.5(5)	5.7×10^{-32}	1(1)
DE436	9.7×10^{-31}	0.8(7)	2.0×10^{-31}	0.9(7)	1.9×10^{-31}	1(1)
BayesEphem	1.3×10^{-31}	0.1(9)	2.7×10^{-32}	0(1)	-4.3×10^{-32}	0(1)

Note. None of the S/Ns are significant.

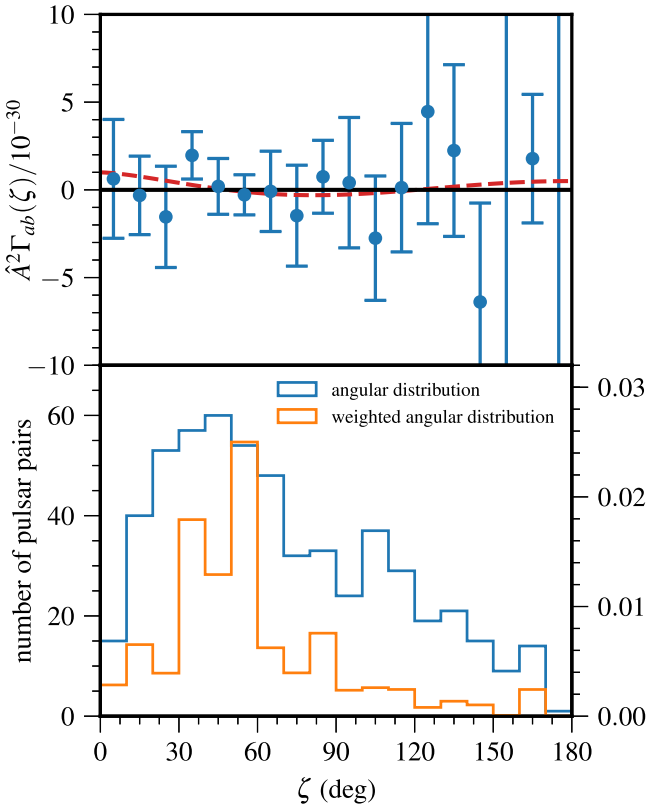


Figure 7. Top: angular distribution of cross-correlated power between pulsar pairs in the NANOGrav 11 year data set, averaged over 10 degree bins. This analysis was done using DE436; other ephemerides give similar results. A GWB would cause the cross-correlated power to lie along the Hellings–Downs curve (red dashed line), shown assuming a GWB amplitude of $A_{\rm GWB} = 10^{-15}$. Bottom: histogram of pulsar-pair angular separations. The blue curve shows numbers in each bin, while the orange curve is reweighted by squared 1σ uncertainties of the averaged cross-correlated power in that bin. Currently, NANOGrav is most sensitive at angular separations between 30° and 60°.

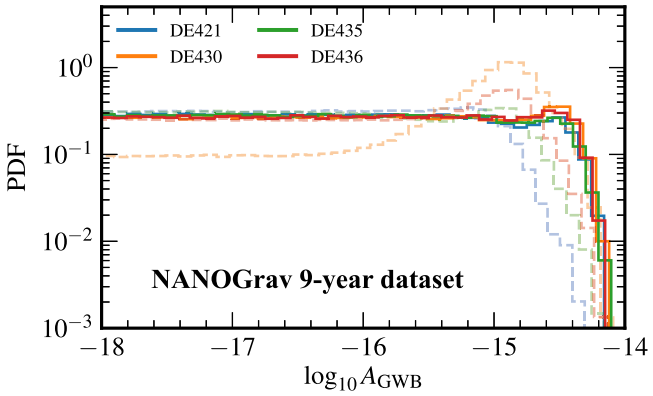


Figure 8. Posterior probability distributions for $A_{\rm GWB}$ (log-uniform prior, $\gamma = 13/3$, and no spatial correlations), as computed for the NANOGrav nine-year data set under individual JPL ephemerides (dashed lines) and with BayesEphem, taking each of the JPL ephemerides as a starting point (solid lines).

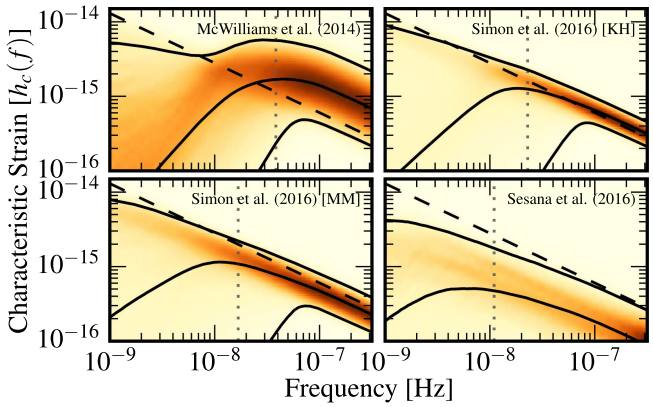


Figure 9. Posterior density plots of broken power-law GWB spectra, as constrained by the 11 year data set, adopting high-frequency $A_{\rm GWB}$ priors from MOP14, Simon & Burke-Spolaor (2016), and S16. At each frequency, a thin vertical segment of the plot should be understood as a density plot of the characteristic strain; the solid lines mark the 2.5%, 50%, and 97.5% quantiles, the dashed line shows the $\gamma = 13/3$ unbroken power-law upper limit, and the vertical dotted lines show the median $f_{\rm bend}$ values. All plots were produced for a spatially uncorrelated common GWB process. Larger values of $A_{\rm GWB}$ induce stronger turnovers at higher frequencies (e.g., MOP14 has a median value (50% quantile) of $f_{\rm bend}$ that is more than three times that of Sesana et al. 2016).

data. This trend is also seen in the increased preference toward larger $\rho_{\rm stars}$ and e_0. While this effect was observed in earlier work, the methodology used in this paper allows for its robust exploration.

Taken at a glance, the results detailed in this work appear less constraining then those presented in NG9b. This is to be expected: marginalizing over parameters, rather than fixing them to set values, will insert more uncertainty into any constraint. Additionally, the methods used in NG9b incorporated an intermediate step by extrapolating from the posterior on $A_{\rm GWB}$ while assuming a power-law GWB. In this paper, we are able to constrain the entire spectrum directly from the data set with no intermediaries—a benefit of the GP method of Taylor et al. (2017b), which will enable future NANOGrav data sets to place constraints on the dynamics of the most massive black holes in the universe.

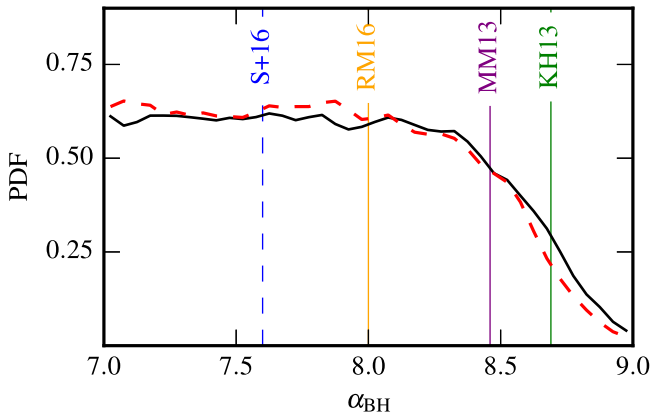


Figure 10. Constraints on $\alpha_{\rm BH}$ from the NANOGrav 11 year data set. The black solid line is the posterior PDF marginalized over the combined parameter space $\{\rho_{\rm stars}, e_0\}$, while the red dashed line is the posterior PDF marginalized over $\rho_{\rm stars}$ for circular binaries ($e_0 = 0$). The red line is slightly more constraining, which is to be expected when a degree of freedom is removed. We do not set a 95% upper limit from these posteriors, since that number would be dependent on the lower bound of the $\alpha_{\rm BH}$ prior. The colored lines show selected observational measurements and predictions for $\alpha_{\rm BH}$: KH13 (Kormendy & Ho 2013), MM13 (McConnell & Ma 2013), RM17 (Rasskazov & Merritt 2017b), S+16 (Shankar et al. 2016). The S+16 line is dashed because that measurement is not a simple power-law relation but includes higher-order terms; here we base our plot on the leading-order coefficient.

5.2. Cosmic Strings

Cosmic strings are linear topological defects that can form in the early universe as a result of symmetry-breaking phase transitions (Kibble 1976; Vilenkin 1981, 1985; Vilenkin & Shellard 2000). Strings that form with lengths greater than the horizon are known as "long" or "infinite" strings, while smaller strings form loops. If two strings meet one another, they can exchange partners, and small portions of string can be chopped off with a reconnection probability p. For classical strings $p = 1$, but string-theory-inspired models may have $p < 1$. This is due to the fact that fundamental strings interact probabilistically, and also that in these models an intersection occurring in the usual three spatial dimensions need not occur in higher compactified dimensions. Cosmic-string networks evolve toward an attractor solution known as the "scaling regime" in which the statistical properties of the system (such as the average size of loops or the distance between long strings) scale with the cosmic time, and the energy density of the string network is a small constant fraction of the radiation or matter density. Cosmic strings have tensions equal to their mass per unit length, μ. This tension is so high that strings oscillate relativistically under their own tension, decaying solely through the emission of GWs, and shrinking in size. The formation of loops and their subsequent decay by GW emission is the mechanism by which the string network loses energy and reaches the scaling regime. The GW spectrum from cosmic-string networks is exceptionally broadband, covering all regions of LIGO, LISA, and PTA sensitivity. For our purposes, we describe the parameter space of cosmic strings in terms of their dimensionless tension, $G\mu/c^2$, and their reconnection probability, p.

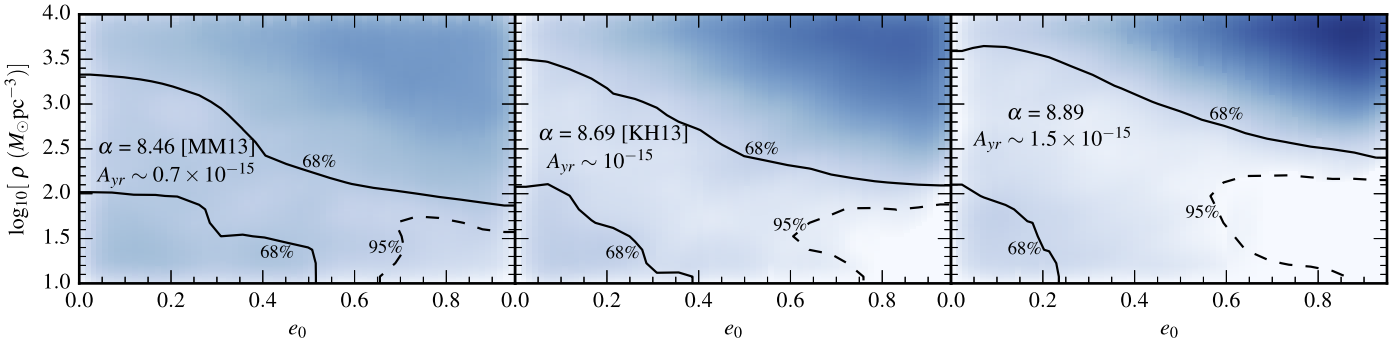
We take a more self-consistent approach than previous PTA analyses. Rather than refit posterior samples (from power-law or free spectrum searches) to cosmic-string models (Lentati et al. 2015; Arzoumanian 2016), we train a GP interpolant on output from the most up-to-date string population simulations. Blanco-Pillado & Olum (2017) and Blanco-Pillado et al. (2017) performed a complete end-to-end calculation of the stochastic GWB expected from a network of cosmic strings, namely (i) simulation of the long-string network to find a representative sample of loop sizes and shapes, (ii) modeling of loop-shape deformations due to gravitational backreaction, (iii) GW spectrum computed for each loop, (iv) evaporation and production modeled to find the distribution of loops over z, (v) integration of the GW spectrum of each loop over the redshift-dependent loop distribution, and finally, (vi) integration over cosmological time to find the present-day GW background.

The output from these simulations corresponds to GW energy-density spectra at a range of string tension values, $G\mu/c^2$, over 25 orders of magnitude in frequency and has been made publicly available.[63] We convert these to characteristic strain, then at each frequency bin in our PTA analysis we train a GP to emulate the strain as a function of string tension. We expand our model to include reconnection probability, p, by analytically scaling the fiducial $p = 1$ strain spectrum by $(1/p)^{1/2}$ (Sakellariadou 2005). We then use this model (with all features of the cosmic-string spectrum included) to analyze the NANOGrav 11 year data set. We do not model signal finiteness or anisotropy due to bright resolvable cosmic-string bursts, since this is only expected when initial loop sizes are very small ($\lesssim 10^{-8}$; Kuroyanagi et al. 2017).

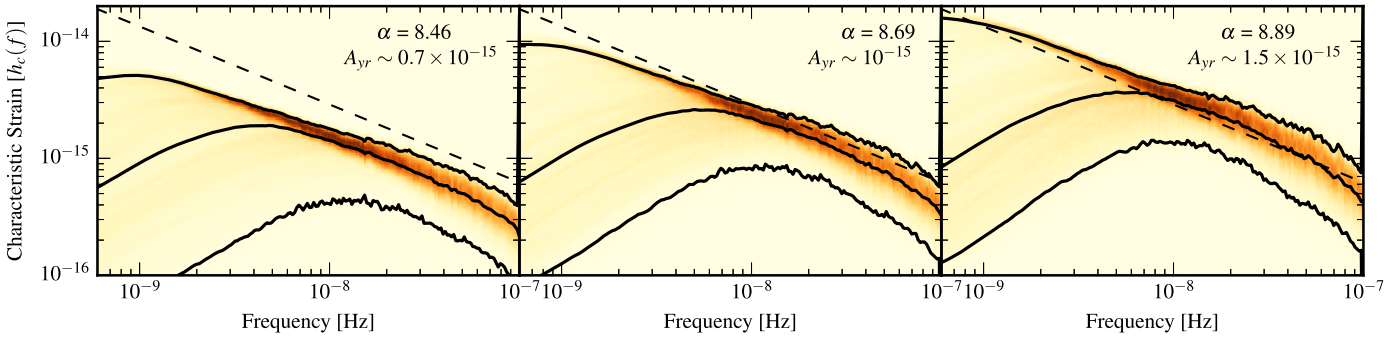
Figure 12 shows the 95% upper limit on string tension as a function of reconnection probability. The shaded region enclosed by the solid black line indicates parameter space that is excluded by the NANOGrav 11 year data set under the assumptions of the Blanco-Pillado & Olum (2017) cosmic-string simulations. For $p = 1$, the string tension is constrained to be $G\mu/c^2 < 5.3(2) \times 10^{-11}$. At this level, we would not expect any measurable effects in the CMB power spectrum, nor through gravitational lensing (Blanco-Pillado et al. 2017). PTAs are currently the best experiment with which to detect cosmic strings and to place stringent limits on the string parameter space.

By contrast, the NANOGrav nine-year data set (NG9b) constraints on string tension (shown as an excluded region with a dashed line boundary) were computed under the assumptions of older string simulations (Blanco-Pillado et al. 2014) and were obtained by resampling the posterior distribution of a power-law GWB spectrum. For $p = 1$, the string tension was constrained to be $G\mu/c^2 < 1.3 \times 10^{-10}$. Finally, even though the most recent EPTA constraints on cosmic strings (Lentati et al. 2015) were not computed under the assumptions of the Blanco-Pillado et al. (2014) simulations, in NG9b the constraints were converted to get a corresponding limit on the string tension of $G\mu/c^2 < 8.6 \times 10^{-10}$. Thus, the constraints on cosmic-string tension from the NANOGrav 11 year data set are 2.5 times better than those from the NANOGrav nine-year data set, and 16.2 times better than those from the most recent EPTA analysis. The 9 to 11 year improvement is to be expected, since BayesEphem analyses of the 11 year data set give consistently more constraining GWB limits than DE421 analyses of the nine-year data set. There are a few other notable caveats to these comparisons: (i) the NANOGrav nine-year and EPTA analyses were performed under a fixed JPL SSE model, while the NANOGrav 11 year

[63] http://cosmos.phy.tufts.edu/cosmic-string-spectra/



(a) 2D posterior for $\rho_{\rm stars}$ and e_0 for fixed $\alpha_{\rm BH}$



(b) Marginalized spectral shape of GWB

Figure 11. Top (a): posteriors for $\rho_{\rm stars}$ and e_0 at different values of $\alpha_{\rm BH}$, as computed for the NANOGrav 11 year data set. Bottom (b): marginalized spectral densities computed from those posteriors. Each column of plots corresponds to a different value of $\alpha_{\rm BH}$ with values decreasing from right to left. The center and left columns correspond to the measured values from Kormendy & Ho (2013) and McConnell & Ma (2013), respectively, while the right column corresponds to a larger value, for comparison. The solid (dashed) line in (a) corresponds to the 68% (95%) contour, and the blue shading is consistent across all of the plots. The dashed and solid lines in (b) are identical to those shown in Figure 9, where the dashed line indicates our upper limit on $A_{\rm GWB}$ of $1.34(1) \times 10^{-15}$ on a power-law GWB ($f^{-2/3}$), and the solid lines show the 2.5%, 50%, and 97.5% confidence levels. As $\alpha_{\rm BH}$ increases, so too does the overall level of the background, and the spectral shape of the GWB is more constrained by the data.

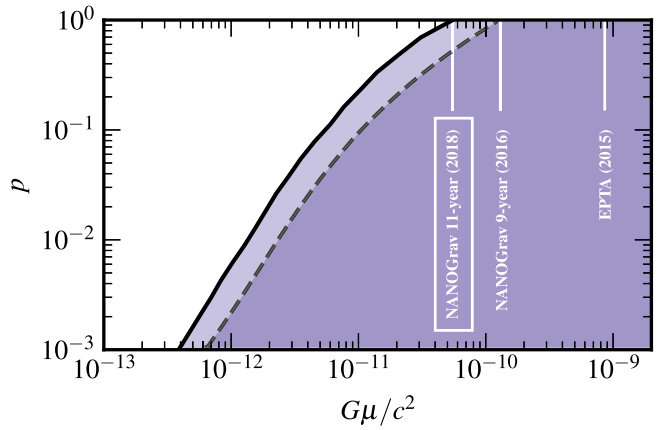


Figure 12. Constraints on cosmic-string tension, $G\mu/c^2$, as a function of reconnection probability, p, with the NANOGrav 11 year data set. The excluded region of parameter space is bounded by a solid black line. The corresponding excluded region for the NANOGrav nine-year data set (NG9b) is bounded by a dashed black line, while the EPTA constraints (Lentati et al. 2015) are shown for $p = 1$ only.

analysis uses BayesEphem, and (ii) the simulation advances of Blanco-Pillado & Olum (2017) with respect to Blanco-Pillado et al. (2014) impede a direct comparison. However, the additional ∼2 years of data in the new NANOGrav data set, the new SSE uncertainty modeling, and the improved end-to-end analysis with simulated cosmic-string spectra all combine to increase NANOGrav's sensitivity to the cosmic-string parameter space.

5.3. Primordial Gravitational Waves

According to the theory of inflation, quantum fluctuations in the spacetime geometry of the early universe are amplified to cosmological scales. Inflation leaves a background of relic primordial GWs that may be observable today (Grishchuk 1976, 1977; Starobinsky 1980; Linde 1982; Fabbri & Pollock 1983). Studies of the cosmic microwave background (CMB) that attempt to observe these GWs indirectly through their imprint of tensor-mode CMB polarizations are limited to probing the surface of last scattering, roughly 300,000 years after the big bang (Kamionkowski et al. 1997; Seljak & Zaldarriaga 1997; BICEP2/Keck et al. 2015). By contrast, GW observations can in principle observe a much earlier epoch in the history of the universe, extending back to as soon as 10^{-32} s post-big bang. Indeed, the spectral index of the primordial GWB is determined by the equation-of-state parameter w in the immediate post-inflation, pre-Big Bang nucleosynthesis universe, and by the tensor index n_t, which depends on the detailed dynamics of inflation (see Grishchuk 2005 and references therein). The primordial spectral dependencies are typically stated in terms of GWB-α as in Lasky et al. (2016) and NG9b. We can express GWB-γ (see Equation (4)) for a primordial

Table 9
NANOGrav 11 Year Upper Limits on Primordial GWs (Last Digit Uncertainty): 95% Credible Intervals Obtained under Uniform Priors for GW and Pulsar Red-noise Amplitudes, and Quoted at Reference Frequency $f_{\rm yr} = {\rm yr}^{-1}$

	95% Upper Limit on $A_{\rm GWB}$ [$\times 10^{-15}$]		
Ephemeris	KE Dom. ($\gamma = 4$)	Rad. Dom. ($\gamma = 5$)	Mat. Dom. ($\gamma = 7$)
DE421	2.01(3)	0.81(1)	0.100(3)
DE430	2.32(2)	0.92(1)	0.117(2)
DE435	2.04(2)	0.84(1)	0.105(2)
DE436	2.10(2)	0.88(1)	0.111(1)
BayesEphem	1.78(2)	0.74(1)	0.099(2)

spectrum as

$$\gamma = \frac{4}{3w+1} + 3 - n_t. \quad (9)$$

In Table 9, we list Bayesian 95% upper limits on $A_{\rm GWB}$ of a primordial GWB, derived as described in Section 3.1. We consider three scenarios, the same considered in NG9b, fixing γ to values corresponding to each: radiation-dominated ($w = 1/3$), matter-dominated ($w = 0$), and kinetic-energy-dominated ($w = 1$) equations of state. Following Zhao (2011), we assume a scale-invariant primordial power spectrum (i.e., $n_t = 0$) for all cases.

These limits constrain the energy-density spectrum of the primordial GWB by way of

$$\Omega_{\rm GWB}(f)h^2 = \frac{2\pi^2}{3{H_0}^2} f^2\ {h_c}^2(f), \quad (10)$$

where h is the dimensionless Hubble parameter, $H_0 = 100\,{\rm km\,s^{-1}\,Mpc^{-1}}$, and h_c is the characteristic GW strain. For a radiation-dominated post-inflationary universe, we obtain

$$\Omega_{\rm GWB}(f_{\rm yr})h^2 \leqslant 3.4(1) \times 10^{-10}, \quad (11)$$

after marginalizing over SSE uncertainties. This is a 20% improvement over the result quoted in NG9b; that number, however, should be significantly revised upward due to SSE bias. Referring back to the bottom panel of Figure 3, we see that the energy-density sensitivity of our PTA data set is dominated by the lowest few frequencies, which individually have 95% upper limit values of $\sim 10^{-9}$, but which in combination beat the limit down to the value quoted in Equation (11).

6. Summary and Conclusions

This paper reports on the search for an isotropic stochastic GWB in NANOGrav's 11 year data set. We targeted a GW signal with predominantly low-frequency power, and so analyzed only those pulsars that have more than three years of observations, corresponding to 34 out of the 45 in the data release. Our investigations encompassed different models of the GWB strain spectrum, spatial correlations between pulsars, and SSE. The latter influence was rigorously studied and led to the major discovery of this paper:

1. We found significant variations in GW upper limits and detection statistics when the data set was analyzed under different published models of the SSE. These models are primarily from JPL, ranging from DE421 to DE436. We also performed a limited analysis with INPOP13c.
2. For a model with H.–D. spatial correlations between pulsars (as appropriate for an isotropic GW background), the 95% upper limit on the amplitude of a fiducial $f^{-2/3}$ power-law strain spectrum (from an astrophysical population of SMBHBs) at a frequency of $1\,{\rm yr}^{-1}$ varies between $1.53–1.79 \times 10^{-15}$.
3. The ratio of Bayesian evidences between models that include a GWB versus only intrinsic pulsar noise processes varies between $\sim$2 and $\sim$26 in favor of a GWB, while the odds favoring GW-induced spatial correlation between pulsars vary between 1.18:1 and 1.63:1. The frequentist analog to the Bayesian odds ratio (known as the "optimal statistic") gives an S/N for GW-induced spatial correlations that varies between 0.57 and 0.87.

This discovery has major ramifications on how we interpret previous PTA results and also how our analysis methodology must be revised for future searches.

1. We formulated a perturbative model ("BayesEphem") that acts to bridge the systematic offsets in the various published models of the SSE, resulting in the first pulsar-timing constraints on GWs that are robust against solar system uncertainties. This model corrects for coordinate-frame drift, uncertainties in gas-giant masses, and uncertainties in Jupiter's orbital elements.
2. Under this new model, the upper limit on the strain amplitude becomes 1.34×10^{-15} for a common red-spectrum process and 1.45×10^{-15} for a GWB. Adding further spatially correlated processes in the model served to worsen these limits only slightly.
3. The evidence ratio for models that include a GWB versus only intrinsic pulsar noise processes is 1 for a GWB with fixed spectral slope and 0.70 if the spectral slope is varied. The odds ratio favoring GW-induced spatial correlations between pulsars is 1.08:1 if the spectral slope is fixed, or 1.15:1 if the slope is varied. The frequentist optimal statistic gives an S/N for GW-induced spatial correlations of 0.09, where the spectral slope is necessarily fixed at the fiducial value of $-2/3$. Both the Bayesian and frequentist analysis show inconclusive evidence for a GW-like red-spectrum process and quadrupolar interpulsar spatial correlations.

We also performed a systematic study of spatially correlated processes in the PTA data set under different ephemerides, tabulating upper limits and evidence ratios for various combinations of a common red-spectrum process, GWB, stochastic clock error, and stochastic SSE uncertainty. With BayesEphem, the presence of these additional spatially correlated processes slightly worsens the GW upper limits, but all remain broadly consistent within uncertainties. Dipole spatial correlations between pulsars seem most disfavored under BayesEphem, likely because we have dealt with the most plausible source of such correlations with our deterministic SSE uncertainty modeling. Uncertainties in the evidence and odds ratios (in addition to their absolute values being around unity) prevent us from being able to make strong

statements. The NANOGrav 11 year data set is only weakly informative of spatial correlations between pulsars.

We used the NANOGrav 11 year data set (with the BayesEphem model) to place constraints on the parameter space of astrophysical and cosmological sources of GWs. As in NG9b, we placed priors on the high-frequency strain amplitude that are motivated by different SMBHB modeling scenarios, then allowed the presence of a turnover in the shape of the strain spectrum to be constrained by the data. With a positive GWB detection, signs of a spectral turnover could indicate that dynamical evolution of SMBHBs remains strongly driven by galactic environmental processes even at centiparsec orbital separations, thereby offering a solution to the "final parsec problem." For a nondetection (as we currently have), this procedure also acts as a test of the validity of our high-frequency strain priors, i.e., priors with larger strain amplitude at high frequency are more in tension with the data when extrapolated back to low frequencies via the fiducial $f^{-2/3}$ scaling, necessitating low-frequency spectral attenuation to ensure consistency with nondetection. We found that the MOP14 prior led to a turnover within the sensitivity band of our PTA ($f > 1/T \sim 3$ nHz) with greater than 97.5% credibility. Other astrophysically motivated priors gave greater consistency with a pure power-law strain spectrum.

For this paper, we took a large step forward in GW spectral modeling and analysis. As described in Taylor et al. (2017b), we trained a GP model on strain spectra from SMBHB population simulations carried out over a large grid in astrophysical parameter space, namely the y-intercept of the $M_{\rm BH}$–$M_{\rm bulge}$ relation ($\alpha_{\rm BH}$), the typical mass density of stars in a galactic core ($\rho_{\rm stars}$), and the binary eccentricity at formation (e_0). This trained model acts as a prior on the GW strain at each frequency, allowing direct recovery of the posterior distribution of astrophysical parameters. We found that the NANOGrav 11 year data set prefers values of $\alpha_{\rm BH}$ that are lower than the largest observed measurements from KH13. Taking the ratios of probability densities as a proxy for model-comparison Bayes factors, we found MM13 to be 1.5 times more probable than KH13, while other, lower measurements are 1.9 times more probable than KH13. These constraints become 2 and 2.6, respectively, when we consider only circular sources (binary eccentricity at formation equaling zero). By studying different values of $\alpha_{\rm BH}$, we showed how larger levels of the GWB, which require spectral shapes that deviate more from the common power law, set progressively tighter constraints on the joint parameter space of $\{\rho_{\rm stars}, e_0\}$. The modeling utilized to produce these results can be trivially expanded to incorporate new astrophysical complexity, and in the era of precision spectral characterization, it will allow PTAs to construct a detailed view of SMBH demographics out to $z \sim 2$.

We took a similar modeling approach for strain spectra resulting from decaying cosmic-string networks, where we calibrated a GP model with the simulations of Blanco-Pillado & Olum (2017). This gave an SSE-marginalized 95% upper limit on the string tension of $G\mu/c^2 = 5.3 \times 10^{-11}$ at a reconnection probability of $p = 1$, which is 2.5 times better than NG9b, and 16.2 times better than Lentati et al. (2015). (These previous published limits were computed without SSE uncertainty modeling, however). PTAs have already surpassed conventional cosmological probes of cosmic-string networks (Lasky et al. 2016) and will continue to offer the best constraints for the foreseeable future. Likewise, we obtained a limit on a background of primordial GWs resulting from the inflation of quantum spacetime fluctuations (with a radiation-dominated post-inflationary universe), corresponding to $\Omega_{\rm GWB} h^2 < 3.4 \times 10^{-10}$ at 95% credibility with SSE marginalization. This is a 20% improvement over NG9b and an even larger improvement once proper SSE modeling is taken into account for the nine-year analysis.

Over the last few years, the PTA community has made great strides in gathering ever larger, higher-quality data sets and in developing sophisticated analysis methods that can deal with the complex noise budgets and subtle systematics typical of pulsar timing, while interfacing ever more closely and robustly with the astrophysics of GW sources. The sequence of recent stochastic-GW papers (for NANOGrav, Demorest et al. 2013, Arzoumanian 2016, and this paper) is a fitting witness to this growth. We expect this effort to be rewarded by nanohertz GW detection within the next several years (Taylor et al. 2016), if the steadfast pursuit of methodological rigor and physical insight remains our cynosure.

We thank the referee for useful suggestions and comments that improved the quality of this manuscript. The NANOGrav project receives support from NSF Physics Frontier Center award number 1430284. NANOGrav research at UBC is supported by an NSERC Discovery Grant and Discovery Accelerator Supplement and by the Canadian Institute for Advanced Research. We thank our colleagues in the International Pulsar Timing Array for comments and useful discussions. We thank Alberto Sesana for commenting on our astrophysical modeling and interpretation. M.V. and J.S. acknowledge support from the JPL RTD program. Portions of this research were carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration. S.R.T. was partially supported by an appointment to the NASA Postdoctoral Program at the Jet Propulsion Laboratory, administered by Oak Ridge Associated Universities through a contract with NASA. S.R.T. thanks E.R.S. for fruitful discussions. J.A.E. was partially supported by NASA through Einstein Fellowship grant PF4-150120. S.B.S. was supported by NSF award #1458952. P.T.B. acknowledges support from the West Virginia University Center for Gravitational Waves and Cosmology. M.A.M. was partially supported by NSF award OIA-1458952. W.W.Z. is supported by the CAS Pioneer Hundred Talents Program and the Strategic Priority Research Program of the Chinese Academy of Sciences grant No. XDB23000000. R.v.H. was supported by NASA Einstein Fellowship grant PF3-140116. This work was supported in part by National Science Foundation grant No. PHYS-1066293 and by the hospitality of the Aspen Center for Physics. Portions of this work performed at NRL are supported by the Chief of Naval Research. This research was performed in part using the Zwicky computer cluster at Caltech supported by NSF under MRI-R2 award No. PHY-0960291 and by the Sherman Fairchild Foundation. A majority of the computational work was performed on the Nemo cluster at UWM supported by NSF grant No. 0923409. Parts of the analysis in this work were carried out on the Nimrod cluster made available by S.M.R. Data for this project were collected using the facilities of the National Radio Astronomy Observatory and the Arecibo Observatory. The National Radio Astronomy Observatory is a facility of the NSF operated under cooperative agreement by Associated Universities, Inc. The Green Bank Observatory is a

facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc. The Arecibo Observatory is operated by SRI International under a cooperative agreement with the NSF (AST-1100968), and in alliance with Ana G. Méndez-Universidad Metropolitana and the Universities Space Research Association. This research is part of the Blue Waters sustained-petascale computing project, which is supported by the National Science Foundation (awards OCI-0725070 and ACI-1238993) and the state of Illinois. Blue Waters is a joint effort of the University of Illinois at Urbana-Champaign and its National Center for Supercomputing Applications. Some of the algorithms used in this article were optimized using the Blue Waters allocation "Accelerating the detection of gravitational waves with GPUs." The Flatiron Institute is supported by the Simons Foundation.

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. Z.A., K.C., P.B.D., M.E.D., T.D., J.A.E., E.C.F., R.D.F., E.F., P.A.G., G.J., M.L.J., M.T.L., L.L., D.R.L., R.S.L., M.A.M., C.N., D.J.N., T.T.P., S.M.R., P.S.R., R.S., I.H.S., K.S., J.K.S., and W.Z. developed the 11 year 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 NG11. S.R.T. coordinated the writing of the paper and led the search. S.R.T., J.A.E., P.T.B., K.P.I., S.J.V., T.T.P., J.S.H., and N.S.P. directly ran the analysis pipelines. S.R.T. and M.V. designed the BayesEphem statistical analysis. S.J.V., K.P.I., and J.A.E. designed and ran the optimal-statistic analysis and interpreted the results. E.H. participated in the optimization of (some) of the gravitational-wave detection pipelines used in this analysis. N.J.C., X.S., M. V., and T.J.W.L. provided feedback on searches and new analysis techniques, as well as vetted the paper in an internal review process. J.S., S.R.T., and M.V. developed the interpretation of the astrophysical results. P.T.B. ran the relic GW analysis and interpreted the results. S.R.T., J.S., and X.S. developed and interpreted the cosmic strings results. S.R.T., J.A.E., J.S., M.V., P.T.B., T.T.P., J.S.H., M.V., X.S., S.J.V, K.P.I, and CMFM wrote the paper, collected the bibliography, and prepared the figures and tables.

ORCID iDs

S. Burke-Spolaor https://orcid.org/0000-0003-4052-7838
S. Chatterjee https://orcid.org/0000-0002-2878-1502
J. M. Cordes https://orcid.org/0000-0002-4049-1882
N. J. Cornish https://orcid.org/0000-0002-7435-0869
F. Crawford https://orcid.org/0000-0002-2578-0360
H. Thankful Cromartie https://orcid.org/0000-0002-6039-692X
M. DeCesar https://orcid.org/0000-0002-2185-1790
T. Dolch https://orcid.org/0000-0001-8885-6388
R. D. Ferdman https://orcid.org/0000-0002-2223-1235
P. A. Gentile https://orcid.org/0000-0001-8158-683X
M. L. Jones https://orcid.org/0000-0001-6607-3710
D. L. Kaplan https://orcid.org/0000-0001-6295-2881
V. M. Kaspi https://orcid.org/0000-0001-9345-0307
M. T. Lam https://orcid.org/0000-0003-0721-651X
L. Levin https://orcid.org/0000-0002-2034-2986
D. R. Lorimer https://orcid.org/0000-0003-1301-966X
R. S. Lynch https://orcid.org/0000-0001-5229-7430
M. A. McLaughlin https://orcid.org/0000-0001-7697-7422
S. T. McWilliams https://orcid.org/0000-0003-2397-8290
C. M. F. Mingarelli https://orcid.org/0000-0002-4307-1322
D. J. Nice https://orcid.org/0000-0002-6709-2566
R. S. Park https://orcid.org/0000-0001-9896-4585
N. S. Pol https://orcid.org/0000-0002-8826-1285
S. M. Ransom https://orcid.org/0000-0001-5799-9714
P. S. Ray https://orcid.org/0000-0002-5297-5278
A. Rasskazov https://orcid.org/0000-0002-4570-8892
J. Simon https://orcid.org/0000-0003-1407-6607
R. Spiewak https://orcid.org/0000-0002-6730-3298
I. H. Stairs https://orcid.org/0000-0001-9784-8670
K. Stovall https://orcid.org/0000-0002-7261-594X
J. Swiggum https://orcid.org/0000-0002-1075-3837
S. R. Taylor https://orcid.org/0000-0003-0264-1453
M. Vallisneri https://orcid.org/0000-0002-4162-0033
W. W. Zhu https://orcid.org/0000-0001-5105-4058

References

Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2016, PhRvL, 116, 061102
Abbott, B. P., Abbott, R., Abbott, T. D., et al. 2017, PhRvL, 119, 161101
Anholm, M., Ballmer, S., Creighton, J. D. E., Price, L. R., & Siemens, X. 2009, PhRvD, 79, 084030
Armitage, P. J., & Natarajan, P. 2005, ApJ, 634, 921
Arzoumanian, Z., Brazier, A., Burke-Spolaor, S., et al. 2014, ApJ, 794, 141
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
Arzoumanian, Z., Brazier, A., Burke-Spolaor, S., et al. 2018, arXiv:1801.01837
Begelman, M. C., Blandford, R. D., & Rees, M. J. 1980, Natur, 287, 307
BICEP2/Keck, Planck Collaborations, et al. 2015, arXiv:1502.00612
Blanco-Pillado, J. J., & Olum, K. D. 2017, arXiv:1709.02693
Blanco-Pillado, J. J., Olum, K. D., & Shlaer, B. 2014, PhRvD, 89, 023512
Blanco-Pillado, J. J., Olum, K. D., & Siemens, X. 2017, arXiv:1709.02434
Bonetti, M., Sesana, A., Barausse, E., & Haardt, F. 2017, arXiv:1709.06095
Brouwer, D., & Clemence, G. M. 1961, Methods of Celestial Mechanics (New York: Academic)
Carlin, B. P., & Chib, S. 1995, Journal of the Royal Statistical Society, Series B (Methodological), 57, 473
Chamberlin, S. J., Creighton, J. D. E., Siemens, X., et al. 2015, PhRvD, 91, 044048
Cornish, N. J., & Littenberg, T. B. 2015, CQGra, 32, 135012
Cornish, N. J., & Sampson, L. 2016, PhRvD, 93, 104047
Cuadra, J., Armitage, P. J., Alexander, R. D., Begelman, M. C., & Alexander, R. D. 2009, MNRAS, 393, 1423
Damour, T., & Vilenkin, A. 2001, PhRvD, 64, 064008
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, The Annals of Mathematical Statistics, 42, 204
Edwards, R. T., Hobbs, G. B., & Manchester, R. N. 2006, MNRAS, 372, 1549
Ellis, J., & van Haasteren, R. 2017a, jellis18/PAL2: PAL2, Zenodo, https://doi.org/10.5281/zenodo.251456
Ellis, J., & van Haasteren, R. 2017b, jellis18/PTMCMCSampler: Official Release, Zenodo, https://doi.org/10.5281/zenodo.1037579
Enoki, M., Nagashima, M., & Nagashima, M. 2007, PThPh, 117, 241
Fabbri, R., & Pollock, M. 1983, PhLB, 125, 445
Fienga, A., Manche, H., Laskar, J., Gastineau, M., & Verma, A. 2014, arXiv:1405.0484
Folkner, W. M., & Park, R. S. 2016, JPL Planetary and Lunar Ephemeris DE436, Tech. Rep., Jet Propulsion Laboratory, Pasadena, CA, https://naif.jpl.nasa.gov/pub/naif/JUNO/kernels/spk/de436s.bsp.lbl

Folkner, W. M., Park, R. S., & Jacobson, R. A. 2016, Planetary ephemeris DE435, Tech. Rep. IOM 392R-16-003, Jet Propulsion Laboratory, Pasadena, CA, ftp://ssd.jpl.nasa.gov/pub/eph/planets/ioms/de435.iom.pdf
Folkner, W. M., Williams, J. G., & Boggs, D. H. 2009, IPNPR, 178, 1 (https://ipnpr.jpl.nasa.gov/progress_report/42-178/178C.pdf)
Folkner, W. M., Williams, J. G., Boggs, D. H., Park, R. S., & Kuchynka, P. 2014, IPNPR, 196, 1 (https://ipnpr.jpl.nasa.gov/progress_report/42-196/196C.pdf)
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
Godsill, S. J. 2001, Journal of Computational and Graphical Statistics, 10, 230
Gregory, P. C. 2005, Bayesian Logical Data Analysis for the Physical Sciences: A Comparative Approach with "Mathematica" Support (Cambridge: Cambridge Univ. Press)
Grishchuk, L. P. 1976, JETPL, 23, 293
Grishchuk, L. P. 1977, NYASA, 302, 439
Grishchuk, L. P. 2005, PhyU, 48, 1235
Gualandris, A., Dotti, M., & Sesana, A. 2012, MNRAS, 420, L38
Haiman, Z., Kocsis, B., & Menou, K. 2009, ApJ, 700, 1952
Hee, S., Handley, W. J., Hobson, M. P., & Lasenby, A. N. 2016, MNRAS, 455, 2461
Hellings, R. W., & Downs, G. S. 1983, ApJL, 265, L39
Hobbs, G. 2013, CQGra, 30, 224007
Hobbs, G. B., Edwards, R. T., & Manchester, R. N. 2006, MNRAS, 369, 655
Hotan, A. W., van Straten, W., & Manchester, R. N. 2004, PASA, 21, 302
Huerta, E. A., McWilliams, S. T., Gair, J. R., & Taylor, S. R. 2015, arXiv:1504.00928
IAU 2017, Division I Working Group on Numerical Standards for Fundamental Astronomy–Astronomical Constants: Current Best Estimates, http://maia.usno.navy.mil/NSFA/NSFA_cbe.html
Ilbert, O., McCracken, H. J., Le Fèvre, O., et al. 2013, A&A, 556, A55
Ivanov, P. B., Papaloizou, J. C. B., & Polnarev, A. G. 1999, MNRAS, 307, 79
Jacobson, R. A. 2009, AJ, 137, 4322
Jacobson, R. A. 2014, AJ, 148, 76
Jacobson, R. A., Antreasian, P. G., Bordi, J. J., et al. 2006, AJ, 132, 2520
Jacobson, R. A., Haw, R. J., McElrath, T. P., & Antreasian, P. G. 2000, JAnSc, 48, 495
Jeffreys, H. 1961, Theory of Probability (3rd ed.; Oxford: Oxford Univ. Press)
Kamionkowski, M., Kosowsky, A., & Stebbins, A. 1997, PhRvL, 78, 2058
Kass, R. E., & Raftery, A. E. 1995, J. Am. Stat. Assoc., 90, 773
Kibble, T. W. B. 1976, JPhA, 9, 1387
Kocsis, B., & Sesana, A. 2011, MNRAS, 411, 1467
Kormendy, J., & Ho, L. C. 2013, ARA&A, 51, 511
Kuroyanagi, S., Takahashi, K., Yonemaru, N., & Kumamoto, H. 2017, PhRvD, 95, 043531
Lasky, P. D., Mingarelli, C. M. F., Smith, T. L., et al. 2016, PhRvX, 6, 011035
Lentati, L., Taylor, S. R., Mingarelli, C. M. F., et al. 2015, MNRAS, 453, 2576
Linde, A. D. 1982, PhLB, 108, 389
McConnell, N. J., & Ma, C.-P. 2013, ApJ, 764, 184
McLaughlin, M. A. 2013, CQGra, 30, 224008
McWilliams, S. T., Ostriker, J. P., & Pretorius, F. 2014, ApJ, 789, 156
Mikkola, S., & Valtonen, M. J. 1992, MNRAS, 259, 115
Mingarelli, C. M. F., & Sidery, T. 2014, PhRvD, 90, 062011
Mingarelli, C. M. F., Sidery, T., Mandel, I., & Vecchio, A. 2013, PhRvD, 88, 062005
Mirza, M. A., Tahir, A., Khan, F. M., et al. 2017, MNRAS, 470, 940
Ölmez, S., Mandic, V., & Siemens, X. 2010, PhRvD, 81, 104028
Peters, P. C. 1964, PhRv, 136, 1224
Peters, P. C., & Mathews, J. 1963, PhRv, 131, 435
Phinney, E. S. 2001, arXiv:astro-ph/0108028
Quinlan, G. D. 1996, NewA, 1, 35
Rasskazov, A., & Merritt, D. 2017a, ApJ, 837, 135
Rasskazov, A., & Merritt, D. 2017b, PhRvD, 95, 084032
Ravi, V., Wyithe, J. S. B., Shannon, R. M., Hobbs, G., & Manchester, R. N. 2014, MNRAS, 442, 56
Robotham, A. S. G., Driver, S. P., Davies, L. J. M., et al. 2014, MNRAS, 444, 3986
Roedig, C., Dotti, M., Sesana, A., Cuadra, J., & Colpi, M. 2011, MNRAS, 415, 3033
Roedig, C., & Sesana, A. 2012, Journal of Physics Conference Series, 363, 012035
Ryu, T., Perna, R., Haiman, Z., Ostriker, J. P., & Stone, N. C. 2018, MNRAS, 473, 3410
Sakellariadou, M. 2005, JCAP, 4, 003
Sampson, L., Cornish, N. J., & McWilliams, S. T. 2015, PhRvD, 91, 084055
Sazhin, M. V. 1978, SvA, 22, 36
Seljak, U. B. U., & Zaldarriaga, M. 1997, PhRvL, 78, 2054
Sesana, A. 2010, ApJ, 719, 851
Sesana, A. 2013, MNRAS, 433, L1
Sesana, A. 2013, CQGra, 30, 224014
Sesana, A., Gualandris, A., & Dotti, M. 2011, MNRAS, 415, L35
Sesana, A., Haardt, F., & Madau, P. 2006, ApJ, 651, 392
Sesana, A., Haardt, F., Madau, P., & Volonteri, M. 2004, ApJ, 611, 623
Sesana, A., Shankar, F., Bernardi, M., & Sheth, R. K. 2016, MNRAS, 463, L6
Shankar, F., Bernardi, M., Sheth, R. K., et al. 2016, MNRAS, arXiv:1603.01276
Shannon, R. M., Ravi, V., Coles, W. A., et al. 2013, Sci, 342, 334
Shannon, R. M., Ravi, V., Lentati, L. T., et al. 2015, arXiv:1509.07320
Sherman, J., & Morrison, W. J. 1950, The Annals of Mathematical Statistics, 21, 124
Siemens, X., Ellis, J., Jenet, F., & Romano, J. D. 2013, CQGra, 30, 224015
Simon, J., & Burke-Spolaor, S. 2016, ApJ, 826, 11
Standish, E. M. 2004, A&A, 417, 1165
Starobinsky, A. A. 1980, PhLB, 91, 99
Taylor, S. R. 2017, https://github.com/stevertaylor/NX01
Taylor, S. R., & Gair, J. R. 2013, PhRvD, 88, 084001
Taylor, S. R., Lentati, L., Babak, S., et al. 2017a, PhRvD, 95, 042002
Taylor, S. R., Simon, J., & Sampson, L. 2017b, PhRvL, 118, 181102
Taylor, S. R., Vallisneri, M., Ellis, J. A., et al. 2016, ApJL, 819, L6
Tiburzi, C., Hobbs, G., Kerr, M., et al. 2016, MNRAS, 455, 4339
Tremmel, M., Governato, F., Volonteri, M., Quinn, T. R., & Pontzen, A. 2018, MNRAS, arXiv:1708.07126
Vallisneri, M. 2012, PhRvD, 86, 082001
Vallisneri, M., & van Haasteren, R. 2017, MNRAS, 466, 4954
van Haasteren, R., Levin, Y., Janssen, G. H., et al. 2011, MNRAS, 414, 3117
Verbiest, J. P. W., Lentati, L., Hobbs, G., et al. 2016, MNRAS, 458, 1267
Vilenkin, A. 1981, PhLB, 107, 47
Vilenkin, A. 1985, PhR, 121, 263
Vilenkin, A., & Shellard, E. P. S. 2000, Cosmic Strings and Other Topological Defects (Cambridge: Cambridge Univ. Press)
Wilcox, R. R. 2016, Introduction to Robust Estimation and Hypothesis Testing (Waltham, MA: Elsevier)
Williams, C. K., & Rasmussen, C. E. 2006, Gaussian Processes for Machine Learning (Cambridge, MA: MIT Press), 4
Zhao, W. 2011, PhRvD, 83, 104021