Harmonic analysis for pulsar timing arrays

Jonathan Nay[,1] Kimberly K. Boddy[,1] Tristan L. Smith[,2] and Chiara M. F. Mingarelli[3,4,5]

[1]*Texas Center for Cosmology and Astroparticle Physics, Weinberg Institute, Department of Physics,*
The University of Texas at Austin, Austin, Texas 78712, USA
[2]*Department of Physics and Astronomy, Swarthmore College, Swarthmore, Pennsylvania 19081, USA*
[3]*Department of Physics, University of Connecticut, 196 Auditorium Road, U-3046, Storrs,*
Connecticut 06269-3046, USA
[4]*Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue,*
New York, New York 10010, USA
[5]*Department of Physics, Yale University, New Haven, Connecticut 06520, USA*

We investigate the use of harmonic analysis techniques to perform measurements of the angular power spectrum on mock pulsar timing data for an isotropic stochastic gravitational-wave background (SGWB) with a dimensionless strain amplitude $A_{\rm gw} = 2 \times 10^{-15}$ and spectral index $\gamma_{\rm gw} = 13/3$. We examine the sensitivity of our harmonic analysis to the number of pulsars (50, 100, and 150) and length of pulsar observation time (10, 20, and 30 years) for an isotropic distribution of pulsars. We account for intrinsic pulsar red noise and use an average value of white noise of ~ 100 ns. We are able to detect the quadrupole for all our mock harmonic analyses, and for 150 pulsars observed for 30 years we are able to detect up to the $\ell = 5$ multipole. We provide linear scaling relationships for the SGWB amplitude, the quadrupole, and $\ell = 3$ as a function of pulsar observation time and number of pulsars. We estimate the sensitivity of our harmonic approach to deviations of general relativity that produce subluminal GW propagation speeds.

I. INTRODUCTION

The direct detection of gravitational waves (GWs) by LIGO [1] has marked the rise of an exciting new observational era. Currently, all reported GW measurements are of individual binary mergers of compact objects, but there are also ongoing efforts to detect a stochastic GW background (SGWB) arising from a population of unresolved binaries.[1] In particular, pulsar timing arrays (PTAs) [5–8] are searching for a SGWB in the nanohertz regime, generated by mergers of supermassive black hole binaries. PTAs measure pulses of radio emission from millisecond pulsars, which serve as precise astronomical clocks due to their highly stable rotational periods. Intervening GWs between Earth and a pulsar (typically $\sim$ a kiloparsec away) induce small shifts in the pulse times of arrival (TOAs), and PTA experiments achieve sensitivity to the effects of GWs by cross-correlating TOA information between pairs of pulsars.

An isotropic SGWB imprints two main signatures in TOA data: low-frequency timing shifts common to *all* observed pulsars (referred to as "common red noise" or a "common process") and an angular correlation of timing shifts between pulsar pairs, known as the Hellings-Downs

(HD) curve [9]. Thus far, multiple PTA collaborations have reported an observed common red noise, broadly consistent with expectations from a SGWB, but have not found evidence of the HD correlation necessary to claim a detection [10–13].[2]

Standard PTA pipelines incorporate a Bayesian analysis to characterize frequency spectrum information and a frequentist analysis [14] to assess the evidence of HD angular correlations (e.g., see Refs. [10–13]). However, an equivalent way to represent the HD angular correlation function is through the angular power spectrum, obtained directly by decomposing the isotropic SGWB into spherical harmonics [15] with multipole ℓ [16]. The angular power spectrum for an isotropic SGWB has a dominant quadrupole ($\ell = 2$) contribution due to the tensorial nature of GWs, while higher multipole contributions scale as $\sim \ell^{-4}$ [16,17].

In this paper, we investigate the capabilities of PTAs to measure the angular power spectrum of an isotropic SGWB. We perform Bayesian analyses with mock PTA data and allow the relative amplitudes of the angular power spectrum multipoles to vary as independent parameters. This flexibility naturally permits a search of an isotropic SGWB that incorporates generic modifications of general

---

[1]There may also be contributions from a cosmological SGWB, sourced by exotic processes in the early Universe [2–4].

[2]However, see the note added.

relativity (GR), which may feature GWs with subliminal speeds and non-Einsteinian polarization modes [16,18–21]. By contrast, standard PTA analyses assume HD correlations, which correspond to an angular power spectrum with specific fixed values for the multipole amplitudes. Measuring the shape of the angular power spectrum may also help readily identify the presence of other correlations, such as monopolar clock errors and dipolar solar system ephemeris errors [11], as well as anisotropies in the SGWB [17,22–24].

In order to assess the ability for PTAs to extract the angular power spectrum, we analyze various mock realizations of PTA data assuming standard GR and an isotropic SGWB with a dimensionless strain amplitude $A_{\mathrm{gw}} = 2 \times 10^{-15}$ and spectral index $\gamma_{\mathrm{gw}} = 13/3$. We consider different numbers of observed pulsars (50, 100, and 150) and lengths of observation time (10, 20, and 30 years) for an isotropic distribution of pulsars. In addition to the overall SGWB amplitude, we treat the relative amplitudes for each multipole from $\ell = 2$ to $\ell = 8$ as independent parameters. We also include intrinsic red and white noise for each pulsar and assume the noise is subdominant (in the low-frequency regime of interest) to the SGWB signal, with an average white-noise value of $\sim$100 ns.

In all our harmonic analyses, we are able to detect the quadrupole. For our harmonic analysis with 150 pulsars observed for 30 years, we are able to detect multipoles up to $\ell = 5$. The first four nonzero multipoles in the angular power spectrum can accurately reconstruct the HD curve due to the sharp dropoff in angular power strength as ℓ increases, so our harmonic analysis approach is a promising tool to help characterize angular correlations present in pulsar timing data.

For the SGWB amplitude, the quadrupole, and $\ell = 3$, we provide linear scaling relationships as a function of experiment observation time and as a function of the number of pulsars in our model. We find that increasing the observation time has a larger scaling effect than increasing the number of pulsars. Longer observation times correspond to accessing lower-frequency GWs, where the GW strain is larger (e.g., for a SGWB arising from supermassive black hole binaries [25]).

Finally, we give an example of how our harmonic analysis method enables us to explore deviations from GR by considering subluminal GW propagation speeds. For our harmonic analysis with 150 pulsars observed for 30 years, we find the constraints vary by multipole and range from $0.95c$ to $0.98c$, where c is the speed of light.

This paper is organized as follows. In Sec. II we review how PTA timing data are modeled and define the angular power spectrum. In Sec. III we present our methods for performing a Bayesian harmonic analysis. We describe how we generate our mock PTA data and show our results in Sec. IV. We conclude in Sec. V. We also include various

supplementary material. In Appendix A we demonstrate our results are not driven by numerical outliers when generating mock data by performing a realization study in which we vary the pseudorandom number generator seeds. In Appendix B we show our method for calculating a Savage-Dickey Bayes factor as a measure of the evidence for each multipole in our model. In Appendix C we provide corner plots showing the marginalized 1D and 2D posterior distributions for all SGWB model parameters of the harmonic analyses presented in Sec. IV.

II. PULSAR TIMING RESIDUALS

PTAs time millisecond pulsars by recording pulses over a window of time of $\lesssim$a few days. Within this window, TOAs are obtained (at multiple radio telescope frequencies) by integrating the pulses over $\sim$1 h in order to increase the signal-to-noise ratio. However, for the purposes of this paper, we consider a wideband approach [26], for which there is a single TOA associated with the full window of time, and we use a single radio telescope frequency, since we are not considering the effects of the dispersion measure on the timing signal. A pulsar timing residual is then obtained by fitting out all known systematic and deterministic processes from the TOA (e.g., see Ref. [27]). PTAs accumulate many timing residuals over the lifetime of the experiment ($\sim$10 yr) by observing a given pulsar with a cadence of 2–3 weeks.

The ith timing residual observed at time t_{ai} from pulsar a located in the $\hat{n}_a$ direction can be expressed as [e.g., see Eq. (23.59) of Ref. [7]]

$$r_a(\hat{n}_a, t_{ai}) = n_a(t_{ai}) + R^{\mathrm{gw}}(\hat{n}_a, t_{ai}) + R^{\mathrm{RN}}_a(t_{ai}) + R^{\mathrm{det}}_{ai}, \quad (1)$$

where $n_a(t_{ai})$ is Gaussian white noise, $R^{\mathrm{gw}}(\hat{n}_a, t_{ai})$ is the contribution from a SGWB, $R^{\mathrm{RN}}_a(t_{ai})$ is the contribution from pulsar intrinsic red noise, and R^{det}_{ai} is any leftover deterministic signal not properly fit out from the TOA [28,29]. The contribution from R^{det}_{ai} is accounted for in our analyses but does not otherwise play an important role for our investigations, so we do not discuss it in detail.

In general, $R^{\mathrm{gw}}(\hat{n}_a, t_{ai})$ also depends on the distance from Earth to the pulsar. However, as we discuss below, we are interested in the two-point correlation function of $R^{\mathrm{gw}}(\hat{n}_a, t_{ai})$ between pairs of pulsars, and the terms involving pulsar distances essentially only contribute for colocated pulsars, doubling the correlation function at vanishing separation (see Sec. 23.3 of Ref. [7] for further discussion). The doubling is commonly referred to as the "pulsar term," which does not depend on the pulsar distances. This approach is consistent with the methodology used by PTA collaborations when searching for a SGWB (e.g., see Ref. [11]) and is justified for actual pulsar distances observed by PTAs [30].

For the remainder of this section, we discuss our modeling of the Gaussian white noise, the red-spectrum

processes stemming from a SGWB and pulsar red noise, and the angular power spectrum of the SGWB, which is the foundation of our harmonic analysis approach.

A. Gaussian white noise

The Gaussian white noise $n_a(t_{ai})$ is taken to have zero mean with a covariance matrix

$$N_{ai,bj} \equiv \langle n_a(t_{ai})n_b(t_{bj}) \rangle = \sigma^2_{\text{WN},ai}\delta_{ab}\delta_{ij}, \qquad (2)$$

where δ_{ab} and δ_{ij} are Kronecker delta functions and the angle brackets denote an ensemble average. The total white-noise variance of the ith TOA measurement for pulsar a is (e.g., see Sec. 3.3 of Ref. [27])

$$\sigma^2_{\text{WN},ai} = S_a^2(\sigma^2_{\text{TOA},ai} + Q_a^2), \qquad (3)$$

where $\sigma_{\text{TOA},ai}$ is the TOA measurement error, S_a is the instrument scaling error (also referred to as EFAC), and Q_a is the instrument quadrature error (EQUAD). We assume a single instrument measures all TOAs for a given pulsar, so there is only one scaling and one quadrature error per pulsar.

Since we assume a wideband approach, in which a single TOA is obtained over a single observation window of a pulsar, the white noise of different TOA measurements should not be correlated (e.g., there is no pulse phase "jitter"). The lack of such a correlated error (ECORR) is represented by the diagonal nature of Eq. (2).

B. Red-spectrum processes

The red-spectrum processes due to a SGWB and pulsar intrinsic red noise are assumed to be stationary and Gaussian (e.g., see Ref. [31]) with zero mean. As with the Gaussian white noise, these processes are thus fully characterized by the two-point correlation functions

$$\langle R^{\text{gw}}(\hat{n}_a, t_{ai})R^{\text{gw}}(\hat{n}_b, t_{bj}) \rangle = \xi^{\text{gw}}(\tau)\zeta(\cos\theta_{ab}), \quad (4)$$

$$\langle R_a^{\text{RN}}(t_{ai})R_b^{\text{RN}}(t_{bj}) \rangle = \xi_a^{\text{RN}}(\tau)\delta_{ab}, \qquad (5)$$

where $\tau \equiv |t_{ai} - t_{bj}|$, $\cos\theta_{ab} = \hat{n}_a \cdot \hat{n}_b$ is the pulsar-pair separation angle, and we have assumed the time dependence and angular dependence of the SGWB correlation function are separable [e.g., see Eqs. (23.60) and (23.64) of Ref. [7]]. Also, since we do not model distances to the pulsars, the condition $\theta_{ab} = 0$ is equivalent to $a = b$. For an isotropic SGWB, the spatial correlation function $\zeta(\cos\theta_{ab})$ is the standard HD curve [9]

$$\zeta(\cos\theta_{ab}) = \frac{1}{2}(1 + \delta_{ab}) - \frac{1}{4}\left(\frac{1 - \cos\theta_{ab}}{2}\right)$$
$$+ \frac{3}{2}\left(\frac{1 - \cos\theta_{ab}}{2}\right)\log\left(\frac{1 - \cos\theta_{ab}}{2}\right), \qquad (6)$$

which has been normalized to 1 for $a = b$. The factor δ_{ab} accounts for an enhancement in the autocorrelation due to the pulsar term [e.g., see Eq. (23.65) of Ref. [7]].

From the Wiener-Khinchin theorem, ξ^{gw} and ξ_a^{RN} each possess a spectral decomposition given by a frequency power spectrum. We parametrize the frequency power spectra for the SGWB and the intrinsic pulsar red noise as power laws of the form [10–13]

$$P^{\text{gw}}(f) \equiv \frac{A_{\text{gw}}^2}{12\pi^2}\left(\frac{f}{f_{\text{yr}}}\right)^{-\gamma_{\text{gw}}}f_{\text{yr}}^{-3}, \qquad (7)$$

$$P^{\text{RN},a}(f) \equiv \frac{A_{\text{RN},a}^2}{12\pi^2}\left(\frac{f}{f_{\text{yr}}}\right)^{-\gamma_{\text{RN},a}}f_{\text{yr}}^{-3}, \qquad (8)$$

where A_{gw} is the dimensionless strain amplitude of the SGWB at a reference frequency $f_{\text{yr}} = 1/\text{yr}$, $A_{\text{RN},a}$ is an equivalent dimensionless amplitude for the intrinsic red noise of pulsar a, and γ_{gw} and $\gamma_{\text{RN},a}$ are the corresponding spectral indices. For a source of inspiraling supermassive black hole binaries, we expect $\gamma_{\text{gw}} \simeq 13/3$ [25]. We assume $\gamma_{\text{RN},a}$ and γ_{gw} are positive, resulting in red spectra, which have more power at lower frequencies. The SGWB spectrum in Eq. (7) is common to all pulsars and thus does not carry the pulsar label a.

We use the same method of modeling the autocorrelation of red-spectrum processes as are used by PTA collaborations to search for a SGWB (e.g., see Ref. [11]). The difference with our harmonic analysis approach is how we model the cross-correlations (i.e., the correlations created from distinct pulsar pairs), which we discuss in the following section.

C. Angular power spectrum

The spatial correlation function $\zeta(\cos\theta_{ab})$ for distinct pulsars a and b can be written as a Legendre polynomial decomposition [17,22,32,33]:

$$\zeta(\cos\theta_{ab}) = \sum_{\ell=2}^{\infty} c_\ell P_\ell(\cos\theta_{ab}), \qquad (9)$$

where $P_\ell(\cos\theta_{ab})$ are Legendre polynomials and c_ℓ are the associated Legendre coefficients. The first nonzero Legendre coefficient is the quadrupole ($\ell = 2$) term. The Legendre coefficients that reconstruct the HD curve for an isotropic SGWB are

$$c_\ell = \frac{3}{2}(2\ell + 1)\frac{(\ell - 2)!}{(\ell + 2)!}, \qquad (10)$$

which exhibit a dominant quadrupolar contribution and a sharp reduction at higher multipoles. The Legendre coefficients in Eq. (10) assume standard GR. Modifications of GR require generalized Legendre coefficients, as shown in

Refs. [16,18–21]. By using a total-angular-momentum basis [15] for the GW metric perturbation, the generalized Legendre coefficients can be written as [16,20]

$$c_\ell \equiv 3(2\ell + 1)|F_\ell|^2, \tag{11}$$

where $|F_\ell|^2$ is the detector response function defined by Eq. (19) of Ref. [16]. From Eqs. (10) and (11), the detector response function under standard GR is

$$|F_\ell|^2 = \frac{1}{2}\frac{(\ell - 2)!}{(\ell + 2)!}. \tag{12}$$

For an isotropic SGWB, the detector response function is related to the angular power spectrum, $C_\ell(\tau)$, by[3] [16,20]

$$C_\ell(\tau) = 12\pi|F_\ell|^2 \times 2 \int_0^\infty df\, P^{\mathrm{gw}}(f) \cos(2\pi f\tau). \tag{13}$$

III. HARMONIC ANALYSIS

In this section, we present our analysis method for measuring the angular power spectrum of an isotropic SGWB. Standard PTA analyses fix the angular correlations of timing residuals to follow the HD curve in Eq. (6). We extend these analysis pipelines by replacing the assumed HD correlation function with the more general form given in Eq. (9) and treating the Legendre coefficients c_ℓ as independent parameters. This generalized approach permits a measurement of the angular power spectrum directly from data, allowing for a more generic search of a SGWB that is agnostic to, e.g., possible modifications of GR.

We incorporate a finite number of multipoles as new parameters, with the general expectation that contributions from higher multipoles are suppressed and can be neglected. As we describe in Sec. IV A, we generate mock data assuming an isotropic SGWB under standard GR, and we find that including coefficients for multipoles (starting at $\ell = 2$) up to $\ell = 8$ is sufficient in our analyses. Note that we could also include the coefficients for $\ell = 0$ and $\ell = 1$, which may arise from clock or solar system ephemeris errors [11], respectively, or from non-Einsteinian polarization modes. Monopole and dipole contributions from non-Einsteinian modes possess their own amplitudes that generally differ from the quadrupolar contribution from GR, while clock and ephemeris errors are expected to have completely different frequency spectra. We leave explorations of monopole and dipole correlations to future work.

For our Bayesian analysis, we consider the likelihood function (e.g., see Ref. [34])

$$p(\vec{r}|\vec{\eta}) = \frac{1}{\sqrt{\det(2\pi\mathcal{C})}} e^{-\frac{1}{2}\vec{r}^T\mathcal{C}^{-1}\vec{r}}, \tag{14}$$

where $\vec{r}$ is a vector of measured timing residuals with entries represented by Eq. (1), $\vec{\eta}$ represents the model parameters of interest, and $\mathcal{C}$ is an analytically marginalized covariance matrix. $\mathcal{C}$ includes white-noise and red-noise covariances in Eqs. (2) and (5), as well as the covariance in Eq. (4) from cross-correlations induced by GWs.

As we describe in Sec. II, the τ-dependent covariance functions in Eqs. (4) and (5) are related to their corresponding frequency power spectra via the Wiener-Khinchin theorem [cf. Eq. (7)]. However, in practice, PTA analyses implement the spectral decomposition using a finite number of sine and cosine basis functions consisting of the lowest harmonics of the fundamental frequency $f_L = 1/T_{\mathrm{obs}}$, where T_{obs} is the total observation time of the experiment [31]. Using the same methodology as PTA collaborations (e.g., see Refs. [10,11]), we use a high-frequency cutoff of $f_{H,\mathrm{gw}} = 14/15\ \mathrm{yr}^{-1}$ for the SGWB and $f_{H,\mathrm{RN}} = 30/15\ \mathrm{yr}^{-1} = 2\ \mathrm{yr}^{-1}$ for the intrinsic red noise.

IV. METHODS AND RESULTS

In order to assess the performance of a harmonic analysis, we want to understand how well the angular power spectrum can be reconstructed under various observational scenarios. In this section, we describe the generation of mock data, outline our analysis pipeline, and present our results.

A. Generating mock data

We create nine different mock PTA datasets with varying observation time T_{obs} (10, 20, and 30 yr) and number of pulsars N_p (50, 100, and 150). Our mock data are generated using a method which is similar to that used in Ref. [35].

We assume all synthetic pulsars have a common observation start time and are observed at the same cadence. We randomly populate 150 pulsars isotropically across the full sky, as shown in Fig. 1. We also show the pulsar locations from the International PTA (IPTA) data release 2 (DR2) [36] for comparison. For the 100 (50) pulsar analyses, we select a random subset of the 150 (100) pulsars. The locations of these pulsars are fixed for all analyses involving the same number of pulsars.

We generate the list of TOAs for each pulsar using TEMPO2 [37] and its PYTHON wrapper libstempo [38]. We assume an average observation cadence of 14 days (with small random variations in time for each individual TOA measurement) to create a set of $T_{\mathrm{obs}}/(14\ \mathrm{days})$ number of TOAs for each pulsar. The TOAs need to be adjusted to account for measurement and instrument error, pulsar intrinsic red noise, and the presence of an isotropic SGWB.

We randomly sample noise parameters from truncated normal, truncated log-normal, or uniform distributions, as

---

[3]In practice, the frequency integral in Eq. (13) employs a lower and upper frequency cutoff, which we provide in Sec. III.
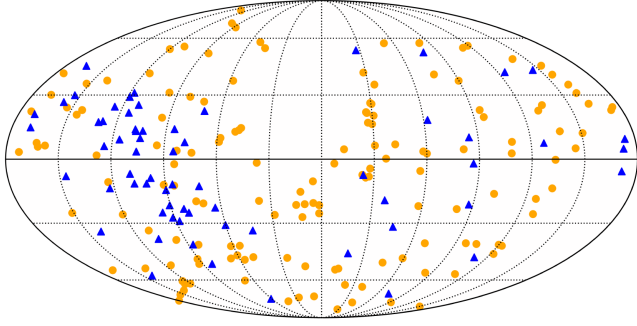
FIG. 1. Mollweide projection of our 150 synthetic pulsars, distributed uniformly across the full sky (orange circles). For comparison, we also show the observed pulsars from IPTA DR2 (blue triangles) [36].

listed in Table I. For each pulsar a, we generate values for instrument errors (S_a and Q_a) and intrinsic red noise ($A_{\mathrm{RN},a}$ and $\gamma_{\mathrm{RN},a}$). Furthermore, the ith TOA receives a measurement error ($\sigma_{\mathrm{TOA},ai}$). Note that we set the maximum on the ranges for $A_{\mathrm{RN},a}$ and $\gamma_{\mathrm{RN},a}$ so that there is a high probability the SGWB signal will dominate the pulsar intrinsic red noise at lower frequencies where the SGWB is modeled. These maximum ranges help ensure each synthetic pulsar contributes to a SGWB signal that is distinguishable from the noise of that pulsar.

We inject a SGWB signal into the mock TOAs, as described in Ref. [40] and we briefly summarize in Appendix A. We use the frequency power spectrum in Eq. (7) with spectral index $\gamma_{\mathrm{gw}} = 13/3$ and amplitude $A_{\mathrm{gw}} = 2 \times 10^{-15}$, which corresponds to the lower end of the reported common red-noise process [10–13].

There are additional pulsar properties we must specify in order to use the full suite of PTA analysis software (i.e., pulsar spin frequency, spin-down rate, parallax, and dispersion measure). We populate the values of these properties by randomly sampling from empirical

distributions that we create using IPTA DR2 pulsar attributes from Ref. [36]. These additional properties do not otherwise play a role in or impact our study, so we do not discuss them further.

Each of our nine mock PTA datasets are generated with the same set of pseudorandom number generator seeds used to inject the white-noise, red-noise, and SGWB signals. We use the same set of seeds in order to focus on differences arising from varying T_{obs} and N_p, without the confounding effects of different realizations of pseudorandom number generation. To ensure the results of our mock datasets are not driven by outlier seeds, we perform a realization study, described in Appendix A. Each realization analysis involves creating mock data with 100 different sets of seeds to produce 100 different realizations, and all other aspects of the analysis remain the same. We find that the main qualitative results of this paper are unchanged with different mock PTA dataset realizations.

B. Analysis methods

In order to implement the harmonic analysis presented in Sec. III, we use ENTERPRISE [41] and ENTERPRISE-EXTENSIONS [42] to calculate the likelihood in Eq. (14). We modify ENTERPRISE-EXTENSIONS to include the angular correlation from Eq. (9) with the Legendre coefficients as model parameters. We use PTMCMCSampler [43] to perform Markov chain Monte Carlo (MCMC) sampling to determine parameter posterior distributions from each of our nine mock datasets. The parameters and their prior ranges are listed in Table II. The lower limit of the Legendre coefficient prior range comes from the requirement that the angular power spectrum given by Eq. (13) is a strictly positive quantity. The upper limit of the Legendre coefficient prior range comes from the requirement that the SGWB two-point correlation function given by Eq. (4) must be positive definite.

TABLE I. Pulsar white- and red-noise parameters for generating mock PTA data. The first column denotes the noise parameter of interest, with units indicated where relevant. In order to randomly generate values of these parameters for each pulsar a and/or TOA i, we sample most of the parameters using a truncated normal distribution with a mean, standard deviation, and range given in the remaining columns. We sample γ_{RN} using a uniform distribution. The values for $\sigma_{\mathrm{TOA},ai}$ correspond to optimistic-quality pulsars in Ref. [39].

Noise input	Mean	Standard deviation	Range
$\sigma_{\mathrm{TOA},ai}$ [ns]	100	30	[9, 601]
S_a	1.0	0.05	[0.5, 5.0]
$\log_{10}(Q_a[\mathrm{s}])$	−8.5	0.1	[−11, −6]
$\log_{10} A_{\mathrm{RN},a}$	−16	1	[−18, −14.7]
$\gamma_{\mathrm{RN},a}$		Uniform distribution	[1, 5]

TABLE II. Uniform prior ranges for MCMC single-pulsar and harmonic analyses. The single-pulsar analysis does not include angular correlations or the frequency power spectrum for a SGWB. The white-noise parameters S_a and $\log_{10} Q_a$ in the harmonic analysis are fixed to their maximum-likelihood values obtained from the single-pulsar analysis.

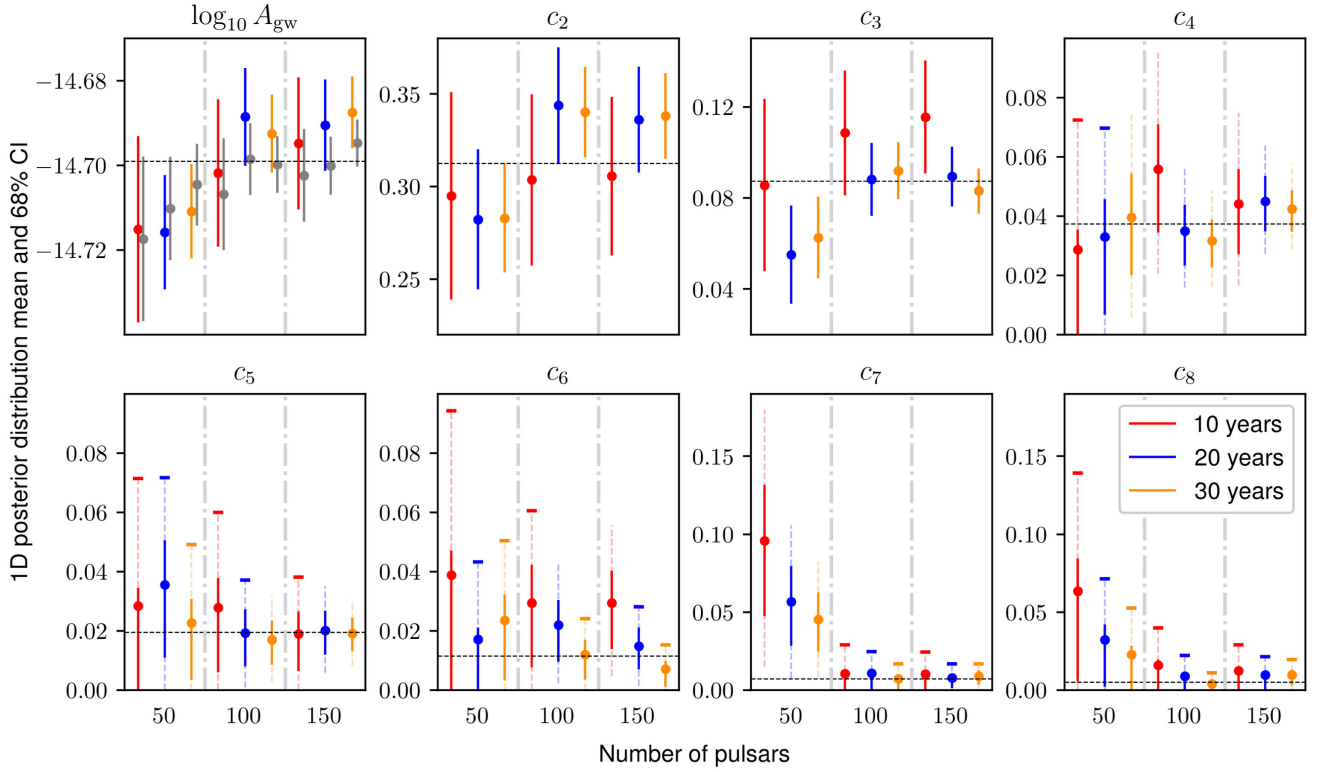| Parameter	MCMC prior range	
	Single-pulsar analysis	Harmonic analysis
$\log_{10} A_{\mathrm{gw}}$	$\cdots$	[−18, −14]
γ_{gw}	$\cdots$	Fixed at 13/3
c_2 through c_8	$\cdots$	[0, 1]
S_a	[0.01, 10]	Fixed to best fit
$\log_{10} Q_a$	[−12, −5]	Fixed to best fit
$\log_{10} A_{\mathrm{RN},a}$	[−20, −11]	[−20, −11]
$\gamma_{\mathrm{RN},a}$	[0, 7]	[0, 7]

FIG. 2. Mean and 68% CI of the marginalized 1D posterior distributions for $\log_{10} A_{\mathrm{gw}}$ and Legendre multipole coefficients for $\ell = 2$ through $\ell = 8$ for all nine harmonic analyses. Each panel separates the analyses for different numbers of pulsars with a vertical dot-dashed line, and analyses for different observation times are shown in different colors, as indicated in the legend. For multipoles $\ell = 4$ through $\ell = 8$, we also show the 95% CI as dashed lines. When the parameter is consistent with zero in the 95% CI, we show a line cap at the upper 95% limit. For each parameter, the injected value is shown as a horizontal dashed, black line. For $\log_{10} A_{\mathrm{gw}}$, the gray points next to each colored point represent the corresponding HD analyses, in which the Legendre coefficients are fixed to their injected values.

Following standard PTA methods [10–13], we perform a single-pulsar noise analysis for a given mock PTA dataset before running the main analysis to extract SGWB properties. The single-pulsar analysis involves four parameters for each pulsar: the two white-noise instrument error parameters (S_a and Q_a) and the two pulsar intrinsic red-noise parameters ($A_{\mathrm{RN},a}$ and $\gamma_{\mathrm{RN},a}$). SGWB parameters are not included: there are no angular correlations for a single pulsar, and the common red-spectrum process of the SGWB cannot be distinguished from the pulsar intrinsic red noise. Thus, the covariance matrix $\mathcal{C}$ in Eq. (14) does not include any cross-correlations between pulsars for the single-pulsar analysis. Note that we use the same number of frequency components to analyze the pulsar intrinsic red noise as those used in the main harmonic analysis.

From the single-pulsar analysis, we obtain maximum-likelihood values for S_a and Q_a. These values for S_a are generally consistent with their injected values (i.e., the values we use to generate the signals for our mock data), while the values for Q_a are poorly recovered, because the injected values of Q_a are subdominant to $\sigma_{\mathrm{TOA},ai}$ [cf. Eq. (3)]. However, in turn, the poorness of the recovery has little impact on our main analysis.

For the full PTA analyses, we fix the values of S_a and Q_a to their maximum-likelihood values obtained from the single-pulsar analysis. The pulsar intrinsic red-noise parameters and the SGWB parameters, including the Legendre coefficients, are varied according to Table II in the MCMC analysis.

We run multiple MCMC chains in parallel to reduce processing time when analyzing a given mock dataset; we do not employ parallel tempering. We combine sampling chains after removing a 25% burn-in to create a single final chain. We use the Gelman-Rubin R statistic [44] as a measure of chain convergence and require $R - 1 < 0.1$ for all SGWB parameters. We also require a minimum combined total of 3.5×10^6 samples prior to removing the burn-in. The sample chains are thinned by a factor of 10, which is the default for PTMCMCSampler. The initial values of the Legendre coefficients, randomly drawn from their prior ranges, are scaled by the total number of Legendre coefficients to prevent MCMC from getting stuck at the initial sample point.

As a point of comparison, we also perform an HD analysis by fixing the values of the Legendre coefficients to their theoretical values in Eq. (10). Fixing the coefficients is

equivalent to using the HD correlation function given by Eq. (6), modulo any effects due to truncating the angular power spectrum at $\ell = 8$. The only SGWB parameter in these HD analyses is $\log_{10} A_{\mathrm{gw}}$, and we use the same prior range given in Table II for the harmonic analysis.

C. Results

For our nine harmonic analyses, we show the mean and 68% credible interval (CI) of the marginalized 1D posterior distributions for the SGWB parameters in Fig. 2. We provide the corner plots for all SGWB parameters of our nine harmonic analyses in Appendix C. For the SGWB amplitude, the quadrupole, and multipole $\ell = 3$, the posterior distributions can be accurately characterized by the mean and 68% CI (as shown in Appendix C). For multipoles $\ell = 4$ through $\ell = 8$, we also show in Fig. 2 the 95% CI with dashed error bars. When the multipole's posterior distribution cannot be distinguished from zero with a 95% CI, we place a line cap at the 95% CI upper limit. The injected value is plotted for each parameter as a dashed black horizontal line. We also show the results of our HD analyses, in which the Legendre coefficients are fixed to their injected values from Eq. (10), for $\log_{10} A_{\mathrm{gw}}$ in Fig. 2. The HD analysis results are plotted in gray, immediately to the right of the analogous harmonic analysis.

For detected parameters (i.e., parameters whose 95% CI does not include zero), Fig. 2 shows the spread of the posterior distribution decreases and the mean tends toward the injected value of the parameter as we increase T_{obs} and N_p, which is intuitively reasonable. For parameters that are not detected, their 95% CI upper limits decrease. The relatively large 68% CI that we see with $\ell = 7$ for the 50 pulsar analyses is realization dependent and within the expected range of fluctuations for parameters that are not detected, as we discuss in Appendix A.

For the 50 pulsar analyses, the posterior of $\log_{10} A_{\mathrm{gw}}$ is biased low relative to the injected value of $\log_{10} A_{\mathrm{gw}}$. The bias comes from modeling pulsar intrinsic red noise, which absorbs some of the power from the injected SGWB signal [45]. The magnitude of this bias decreases as we increase T_{obs} and N_p, because the SGWB amplitude in the cross-correlations of the timing data helps to distinguish the SGWB signal from the pulsar intrinsic red noise.

Figure 2 shows the 68% credible interval of the posterior distribution for $\log_{10} A_{\mathrm{gw}}$ has approximately the same scale between the harmonic analysis and the corresponding HD analysis, which means adding multipoles in our harmonic analysis has a minimal affect on our ability to recover the SGWB amplitude. However, we observe in Fig. 2 that the mean of the posterior distribution for $\log_{10} A_{\mathrm{gw}}$ in the harmonic analysis trends toward larger values than the corresponding HD analysis as we increase T_{obs} and N_p.
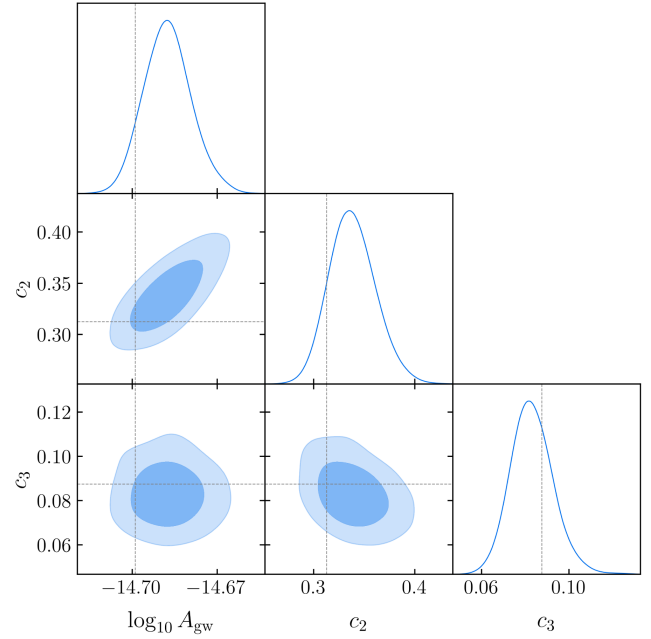

FIG. 3. Corner plot of the marginalized 1D and 2D posterior distributions for $\log_{10} A_{\mathrm{gw}}$, c_2, and c_3 from the harmonic analysis with 150 pulsars observed for 30 yr. The vertical and horizontal dashed lines indicate the injected values. There is a moderately strong positive correlation between $\log_{10} A_{\mathrm{gw}}$ and c_2 and a weak negative correlation between c_2 and c_3. The corner plots for all SGWB parameters in all nine harmonic analyses are provided in Appendix C.

This effect is caused by a moderately strong positive correlation between $\log_{10} A_{\mathrm{gw}}$ and c_2, which we show in the corner plot of Fig. 3. We can see in Fig. 2 that c_2 is trending high relative to its injected value as we increase T_{obs} and N_p, so the positive correlation is causing a similar trend in $\log_{10} A_{\mathrm{gw}}$ for the harmonic analyses. We show in Appendix A that detected parameters generally recover the injected value within 1σ, with fluctuations that are realization dependent and consistent with an ergodic process. Therefore, the trend we observe for $\log_{10} A_{\mathrm{gw}}$ and c_2 is realization dependent.

In Fig. 3 we also observe a weak negative correlation between c_2 and c_3. The weak negative correlation between multipoles is not due to the harmonic mode coupling from a finite number of pulsars [46], because the correlations get stronger as the number of pulsars increases, as can be seen in the corner plots of Appendix C. We expect to observe some correlations between SGWB parameters because of their functional relationship in the angular power spectrum.

In Fig. 4 we plot $\rho_x \equiv \langle x \rangle / \sigma_x$, where $\langle x \rangle$ is the mean of the marginalized 1D posterior distribution and σ_x is one-half the width of the 68% credible interval for $x \in \{A_{\mathrm{gw}}^2, c_2, c_3\}$. Our calculated parameter ρ_x is a measure of the significance of parameter x relative to its uncertainty from its posterior distribution. We obtain the distribution for
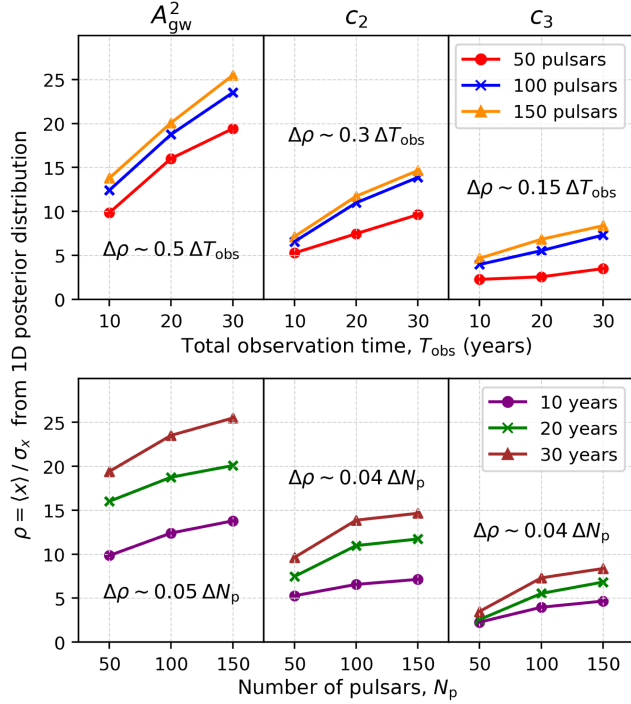
FIG. 4. Ratio of the mean $\langle x \rangle$ to one-half the width of the 68% credible interval σ_x for the marginalized 1D posterior distribution of the parameter $x \in \{A_{gw}^2, c_2, c_3\}$, indicated at the top of each column of panels, for all nine harmonic analyses. We plot the ratios in the top and bottom rows as a function of observation time and number of pulsars, respectively.

FIG. 5. Savage-Dickey Bayes factors $\mathcal{B}_\ell$ for $\ell = 2$ to $\ell = 5$ for all nine harmonic analyses. The Bayes factor provides a measure of the evidence for including a Legendre coefficient c_ℓ in the harmonic analysis. The horizontal dashed lines show the commonly used separation of the Bayes factor strength of evidence: $\log_{10} \mathcal{B}_\ell > 2$ is decisive, $1 < \log_{10} \mathcal{B}_\ell < 2$ is strong, $0.5 < \log_{10} \mathcal{B}_\ell < 1$ is substantial, and $\log_{10} \mathcal{B}_\ell < 0.5$ is no evidence. Coefficients for $\ell > 5$ are not included, since we generally observe $\log_{10} \mathcal{B}_\ell \ll 0.5$ for these multipoles. For well-measured coefficients, the Bayes factor is very large, which we denote as $\log_{10} \mathcal{B}_\ell \gg 5$. We include a small horizontal offset for a few points to prevent overlap.

A_{gw}^2 by transforming the posterior distribution of $\log_{10} A_{gw}$. We choose these three parameters to calculate ρ_x because they are detected in all analyses, as shown in Fig. 2.

The top row of Fig. 4 shows the nine harmonic analyses plotted as a function of T_{obs}, while the bottom row shows these same nine analyses plotted as a function of N_p. We include in Fig. 4 the linear scaling relationship between a change in ρ_x versus a change in T_{obs} (top row) or a change in N_p (bottom row), which we obtain by linear regression. The scaling of A_{gw} with both T_{obs} and N_p roughly matches the scaling shown in Figs. 9 and 11, respectively, of Ref. [47] once we rescale $\rho_{A_{gw}^2}$ by a factor of $1/2$ to account for the fact that we are considering A_{gw}^2, whereas Ref. [47] considers A_{gw}. The difference is largest for the scaling with T_{obs}. This difference may be due to the fact our PTA properties differ and that we perform a full MCMC, whereas Ref. [47] fixes all other parameters to their maximum-likelihood value.

Increasing T_{obs} has a larger affect on ρ_x than increasing N_p. This effect occurs, because increasing T_{obs} increases the number of low-frequency harmonics where the strength of the SGWB signal in the autocorrelations dominates the white noise [48,49]. For our Bayesian analysis approach, we model a finite number of harmonics of $f_L = 1/T_{obs}$, as
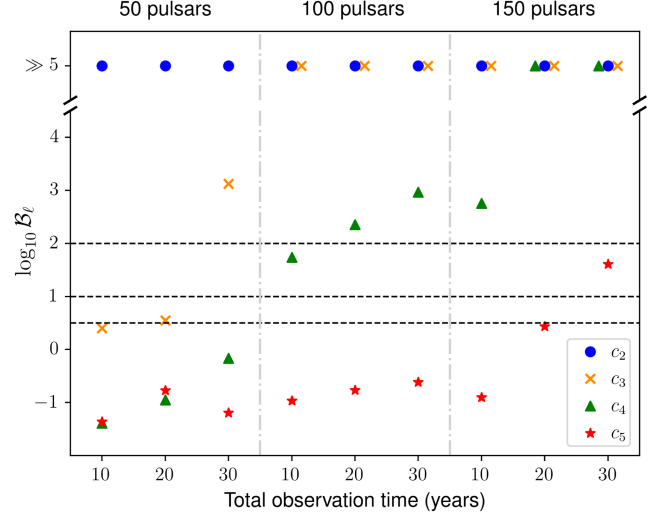
discussed in Sec. III. By using the average value of our injected white noise, we find for $T_{obs} = 10$ yr only the first few harmonics of f_L are in the regime where the SGWB signal is dominant, while for $T_{obs} = 30$ yr the first ten frequency harmonics are in this regime. So increasing T_{obs} provides more frequency bins where the strength of the SGWB signal dominates the white noise in the autocorrelations, thereby reducing the spread of the SGWB amplitude distribution and increasing $\rho_{A_{gw}^2}$.

We use the Savage-Dickey approach [50] to calculate a Bayes factor for each multipole in our harmonic analyses, which is a measure of the evidence for the multipole in our model. The calculation methodology is provided in Appendix B. Figure 5 shows the Savage-Dickey Bayes factors for multipoles up to $\ell = 5$ in the nine harmonic analyses. General trends in Fig. 5 are consistent with basic expectations; i.e., evidence for including higher multipoles increases as T_{obs} and N_p increase due to increasing $\rho_{A_{gw}^2}$. We see there is decisive evidence for the quadrupole in all harmonic analyses. As $\rho_{A_{gw}^2}$ increases, evidence for $\ell = 3$ and $\ell = 4$ becomes decisive; for 150 pulsars observed for 30 years, we see strong evidence for $\ell = 5$. Multipoles $\ell > 5$ are not shown in Fig. 5, because they do not show trending evidence (relative to increasing T_{obs} and N_p) for being included in the model.
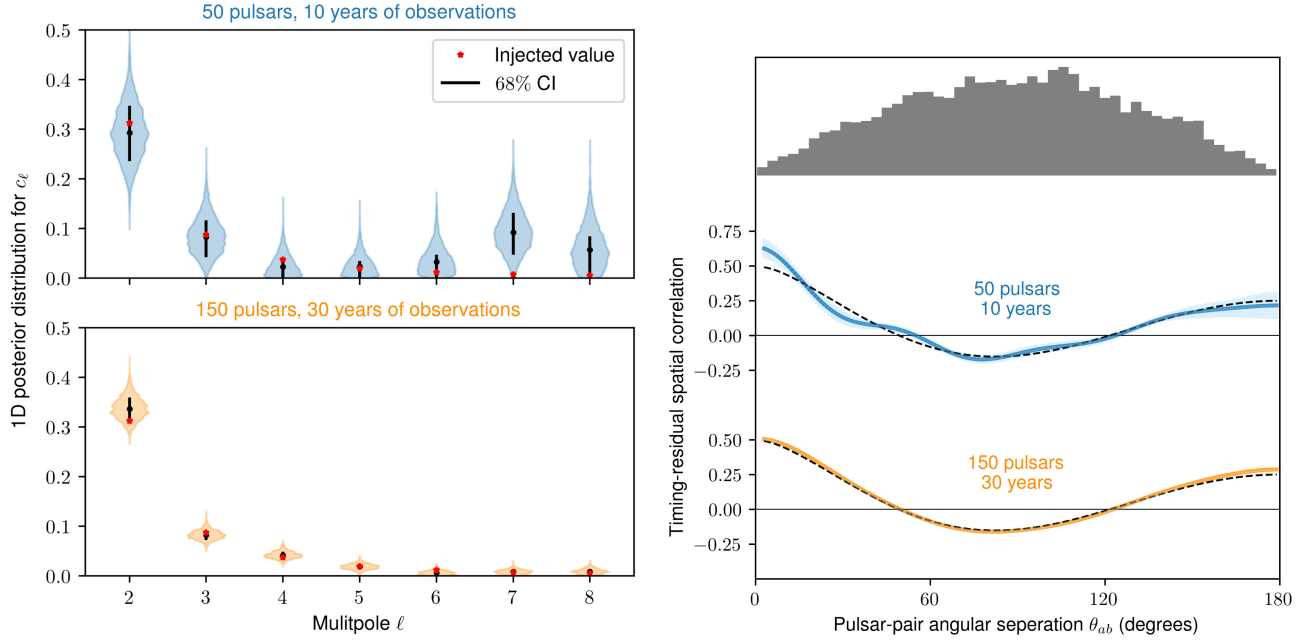
FIG. 6. Left: violin plots of the marginalized 1D posterior distributions for the Legendre coefficients from the two harmonic analyses with the lowest (top panel) and highest (bottom panel) overall multipole evidence. We show the mean (black dot) and 68% CI (black line) for each posterior distribution, along with the injected values of the Legendre coefficients (red stars) from Eq. (10), corresponding the HD angular correlation function. Right: reconstructed angular correlation function using the multipole coefficients shown in the left panels, using 100 angular-separation bins. The solid blue and orange curves represent the mean value within each angular-separation bin, and the error bars show the 1σ deviation of each bin. The dashed black line is the HD curve. At the top of this panel, we show the histogram of angular separations for the 150 synthetic pulsars we use for our harmonic analyses.

The left plot in Fig. 6 shows the multipole marginalized 1D posterior distributions from the harmonic analyses that have the lowest (top) and highest (bottom) overall multipole evidence. These violin plots show the improvement in both the mean (relative to the injected value) and the spread of the posterior distributions as we increase $T_{\rm obs}$ and N_p. The right plot in Fig. 6 shows the reconstructed angular correlation function from the multipoles shown in the left plot. The histogram at the top shows the number of pulsar pairs binned by angular separation. To reconstruct the angular correlation function, the values of the Legendre coefficients at each MCMC chain step are inserted into Eq. (9) at 100 different angular separation bins. The solid blue and orange curves represent the mean value within each angular separation bin, and error bars are the 1σ deviations. The top blue plot shows that even when multipoles $\ell > 3$ have little-to-no evidence for being included in the model, the HD angular correlation function can be fairly accurately reconstructed because of the strong quadrupolar dependency and the sharp dropoff in the angular power spectrum as ℓ increases. The bottom orange plot show how accurately the HD curve can be reconstructed when we have evidence for multipoles up to $\ell = 5$.

Lastly, we provide an example of how our harmonic analysis formalism can be used to constrain scenarios that modify GR. We use the harmonic analysis with 150 pulsars and 30 years of observations to determine the sensitivity to

subluminal GW propagation. The tensor-mode detector response function F_ℓ from Table I of Ref. [20] is a function of the GW subluminal group velocity $v_{\rm gw}$, where we assume a single GW phase velocity with no frequency dependence [to ensure the factorization in Eq. (4) holds]. From Eq. (11), we find $c_{\ell+1}/c_\ell$ as a function of $v_{\rm gw}$, where our use of ratios is to ensure scaling of the autocorrelation cancels. The 95% lower limit of the 1D posterior distributions of c_3/c_2, c_4/c_3, and c_5/c_4 from our harmonic analysis provides corresponding 95% lower limits on the GW group velocity of $v_{\rm gw}/c = 0.95$, $v_{\rm gw}/c = 0.98$, and $v_{\rm gw}/c = 0.96$, respectively. It is interesting to note that, as the group velocity decreases the power spectrum at $\ell > 2$ decreases so that even though the best-measured multipoles are the quadrupole and octupole, our strongest constraint on $v_{\rm gw}$ comes from the ratio of the hexadecapole to the octupole. Note that even though ground-based measurements of GW place tight constraints on the speed of propagation of GWs [1,51], this speed may be different in the frequency range accessible to PTAs [52].

V. CONCLUSIONS

In this paper we demonstrate the capabilities and limitations of a harmonic analysis for mock PTA timing data that includes an isotropic SGWB. We model the first seven nonzero multipoles (i.e., $\ell = 2$ through $\ell = 8$) of the

SGWB angular power spectrum using a Legendre power series representation of the angular correlation function. We perform our harmonic analyses on mock PTA datasets with different numbers of pulsars (50, 100, and 150) and pulsar observation times (10, 20, and 30 years) for pulsars uniformly distributed across the sky.

We find decisive evidence of the quadrupole contribution in the angular power spectrum for all of our harmonic analyses. For higher multipoles, the general trend we observe is that for 50 pulsars evidence for $\ell = 3$ is strong to decisive, for 100 pulsars evidence for $\ell = 3$ is decisive and $\ell = 4$ is strong to decisive, and for 150 pulsars evidence for multipoles up to $\ell = 4$ is decisive and the evidence for $\ell = 5$ becomes strong when evidence of the SGWB amplitude is sufficiently high. The first four non-zero multipoles in the angular power spectrum can accurately reconstruct the HD curve (to within 2% of the spatial correlations integrated over angular separations) due to the sharp dropoff in multipole values as ℓ increases. Therefore, our harmonic analysis approach is a promising tool to help determine angular correlations present in pulsar timing data which includes searching for anisotropies in the SGWB as well as searching for other cosmological and astrophysical sources of GWs.

For A_{gw}^2, c_2, and c_3 we provide linear scaling relationships as a function of T_{obs} and as a function of N_p. We find that increasing T_{obs} has a larger affect than increasing N_p. We also compare the scaling of A_{gw}^2 to previous work and find that our scaling with N_p is consistent with previous work, but our scaling with T_{obs} is higher by approximately a factor of 2, which we suspect is due to different modeling techniques.

For higher multipoles that are not detected in our model, we can place upper limits that decrease sharply with increasing T_{obs} and N_p. We observe that multipoles are not detected until the theoretical value of the Legendre coefficient for that multipole is less than the posterior standard deviation of $\log_{10} A_{\mathrm{gw}}$.

Most analyses of PTA data assume that the angular correlations between pulsars agree with the standard expectations of an isotropic SGWB and GR. As the data improve, it will be essential to allow for us to test these assumptions and allow deviations from these standard expectations. The method we present here provides a flexible parametrization that enables us to more fully explore the implications of any future detection of angular correlations in a PTA, giving us the ability to explore deviations from GR (e.g., Ref. [53]), anisotropy (e.g., Ref. [22]), and systematic errors which cause angular correlations, such as clock errors and shifts in the Solar System barycenter [54].

Note added. Recently, PTA collaborations reported evidence of SGWB spatial correlations consistent with HD correlations at varying levels of significance [55–58].

APPENDIX A: REALIZATION STUDY

In this appendix we present the methodology and results of our realization study in which we vary the pseudorandom number generator seeds. We choose three different mock PTA datasets to perform these analyses: 50 pulsars observed for 10 years, 50 pulsars observed for 20 years, and 100 pulsars observed for 20 years. We choose these datasets so that we can observe the differences in realizations on datasets that have different values of T_{obs} and N_p. In the following subsections, we discuss the methodology used to perform the realization study, provide the realization study results, and compare the realization variance of the quadrupole to predicted variance from Refs. [33,60,61].

1. Realization study methodology

We generate 100 realizations for each of our three selected mock PTA datasets by varying the pseudorandom number generator seeds used to inject white noise, pulsar intrinsic red noise, and the SGWB signal into the timing data. All other analysis methods are the same as those discussed in Secs. III and IV. We provide a brief summary of how pseudorandom number generator seeds inject signals into the timing data using TEMPO2 and its PYTHON wrapper libstempo.

For white noise, we first apply TOA measurement error $\sigma_{\mathrm{TOA},ai}$ to the ith TOA for pulsar a, as we discuss in Sec. IV A. Instrument errors S_a and Q_a, which are the same for all realization analyses of a given mock PTA dataset, adjust the total white noise using Eq. (3). To make these injected white-noise signals Gaussian, S_a and Q_a are scaled by random numbers generated from a standard normal distribution based on specified pseudorandom number generator seeds. We use two different pseudorandom number generator seeds to inject the white noise: one

generator seed for the scaling of S_a and one generator seed for the scaling of Q_a.

We inject intrinsic red noise for pulsar a using the power-law model given by Eq. (8) with values of $A_{\mathrm{RN},a}$ and $\gamma_{\mathrm{RN},a}$ which are the same for all realization analyses of a given mock PTA dataset. The intrinsic red-noise signal is created in the time domain and added directly to the TOAs. The first ten harmonics of $1/T_{\mathrm{obs}}$ are used in a sine-cosine basis representation of the intrinsic red-noise timing signal, as described in Sec. III. The amplitudes of the 20 basis functions (ten sine functions and ten cosine functions) are independently scaled by random numbers generated from a standard normal to make the signal Gaussian and then added to the pulsar TOA. We use a single pseudoran-dom number generator seed for the scaling applied to the time-domain sine-cosine basis function amplitudes of the intrinsic red noise.

The SGWB signal is injected in the frequency domain using the Fourier transform of Eq. (4) and performing a Cholesky decomposition of the angular correlation func-tion, as described in Ref. [40]. The Cholesky decomposi-tion provides a matrix M with elements satisfying $(MM^{\dagger})_{ab} = \zeta(\cos\theta_{ab})$, where $\zeta(\cos\theta_{ab})$ is the HD angular correlation from Eq. (6), which includes the doubling of the autocorrelation via the "pulsar" term. Note that fixing the injected angular correlation to Eq. (6) is equivalent to fixed injected values of each c_ℓ, given by Eq. (10), for all realizations. The SGWB-induced timing residual for pulsar a can then be expressed in the frequency domain as

$$\tilde{R}^{\mathrm{gw}}(\hat{n}_a, f) = c(f)\sum_b M_{ab} w_b(f), \qquad (A1)$$

where $c(f)$ is a frequency-dependent normalization chosen so that the two-point correlation of $\tilde{R}^{\mathrm{gw}}(\hat{n}_a, f)$ gives the SGWB power spectrum from Eq. (7) and $w_b(f)$ is a complex (two-parameter) Gaussian variable for pulsar b at frequency f. In practice, only a finite number of frequency components are modeled; the frequency components range from 0 to the Nyquist sampling frequency based on a 14-day cadence, in increments of $1/(10 T_{\mathrm{obs}})$. For each frequency component, the real and imaginary parts of $w_b(f)$ are independently sampled from a standard normal distribution so that the two-point correlation is given by

$$\langle w_a^*(f) w_b(f') \rangle = \frac{2}{T_{\mathrm{obs}}} \delta(f - f') \delta_{ab}. \qquad (A2)$$

The SGWB-induced timing residual $\tilde{R}^{\mathrm{gw}}(\hat{n}_a, f)$ is trans-formed back into the time domain, and interpolation is used to find the value of $R^{\mathrm{gw}}(\hat{n}_a, t_{ai})$ at observation time t_{ai}, which is then added to the ith TOA for pulsar a.

To summarize, four different pseudorandom number generator seeds are specified for each mock PTA dataset

realization. To perform realization analyses, we vary the set of four generator seeds while ensuring the four generator seeds are all different within a realization.

2. Realization study results

Figure 7 provides a comparison of the multipole evidence from the realization analyses in this appendix against the single realization in Sec. IV C. We calculate the multipole evidence using the Savage-Dickey Bayes factor as discussed in Appendix B. We see in Fig. 7 that the single realization in Sec. IV C has either the same evidence or conservatively low evidence, relative to range of potential evidence from the realization analyses.

In Fig. 8, we plot the mean and 68% CI of parameters $\log_{10} A_{\mathrm{gw}}$, c_2, c_3, and c_4 for each individual realization. We see that the average of the means tends toward the injected value with increasing T_{obs} and N_p. We can also see the mean-value fluctuation scale decreases with increasing T_{obs} and N_p, and we observe that the scale of the mean fluctuation 1σ error bars (shown as black dashed lines) is the same scale as the 68% CI (shown as error bars around each plotted point). In other words, the scale associated with the spread of a detected parameter's posterior 68% CI, which is realization independent, is the same scale as the fluctuation of the parameter's posterior mean, which is realization dependent. Our result is consistent with an ergodic process for SGWB signal fluctuations and explains an additional observation we make with the quadrupole
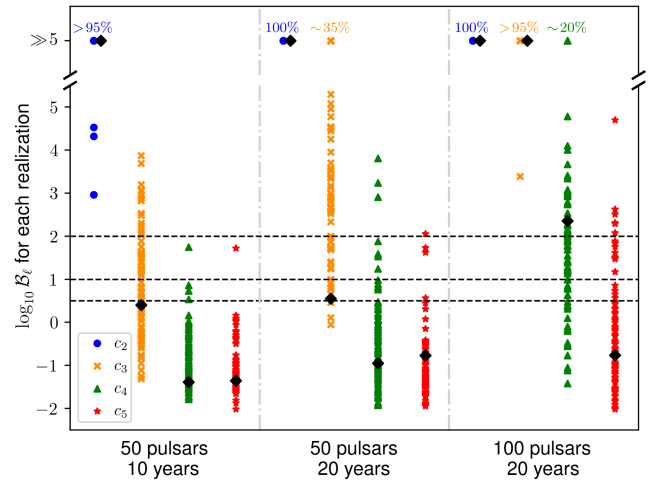

FIG. 7. Savage-Dickey Bayes factors for 100 realizations of three mock PTA datasets: 50 pulsars observed over 10 years, 50 pulsars observed over 20 years, and 100 pulsars observed over 20 years. The Bayes factors for the realization presented in Sec. IV C are plotted as black diamonds. For cases in which a multipole coefficient has a very large Bayes factor for multiple realizations, we indicate the percentage of realizations with $\log_{10}\mathcal{B} \gg 5$. See Fig. 5 for a further explanation of the plot features.
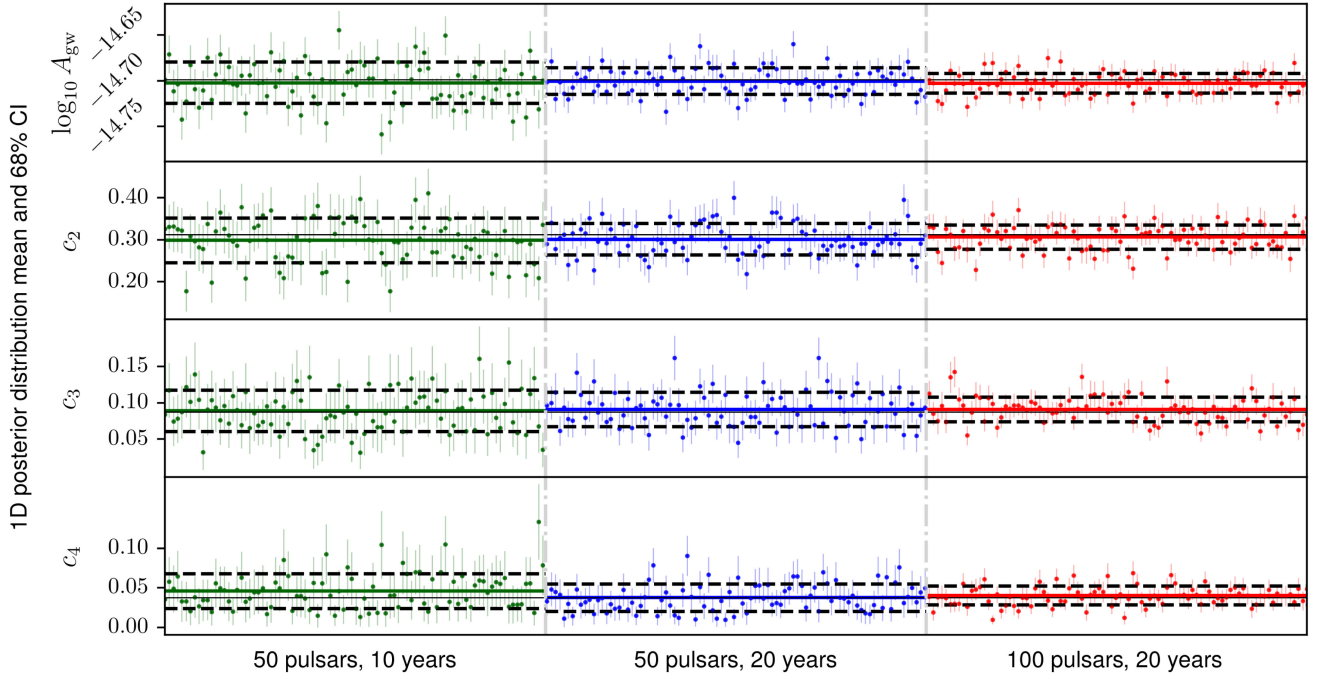
FIG. 8.   Realization-dependent mean and 68% CI for parameters $\log_{10} A_{\mathrm{gw}}$, c_2, c_3, and c_4. Individual realizations are plotted along the x axis. The datasets used for the realization analyses are separated by color and by the gray vertical dot-dashed lines. The colored horizontal solid lines represent the averaged mean of the 1D posterior distributions for all realizations of a given dataset, and the black horizontal dashed lines show the standard deviation of the realization means. The injected value of each parameter is plotted as a solid horizontal black line.

posterior distributions from the realization analyses: with all three datasets, the 68% CI for the quadrupole posterior distribution contains the injected value of the quadrupole approximately 68% of the time.

A final result to mention from the realization analyses is that the realization-averaged standard deviations of the marginalized 1D posterior distributions for multipoles that are not detected are all the same for a given dataset: approximately 0.021, 0.014, and 0.012 for 50 pulsars with 10 years, 50 pulsars with 20 years, and 100 pulsars with 20 years, respectively. These values coincide with the standard deviations of the posterior distribution for $\log_{10} A_{\mathrm{gw}}$ in each dataset, which means $\sigma_{\log_{10} A_{\mathrm{gw}}}$ is related to an inherent harmonic analysis uncertainty for multipoles that are not detected. The parameter $\log_{10} A_{\mathrm{gw}}$ is detected in all analyses, so the realization-dependent fluctuations of $\sigma_{\log_{10} A_{\mathrm{gw}}}$ are small; hence, we can accurately predict this inherent uncertainty with a single realization. Since $c_{6,\mathrm{injected}} \approx 0.0116$ is less than $\sigma_{\log_{10} A_{\mathrm{gw}}}$ for these analyses, we can see why multipoles $\ell \geq 6$ are not detected in these realization analyses. The combination of the spread for nondetected multipoles being comparable to $\sigma_{\log_{10} A_{\mathrm{gw}}}$, along with an ergodic process for the realization-dependent fluctuations, explains why we occasionally see large values for undetected multipoles, such as $\ell = 7$ for the 50 pulsar analyses in Sec. IV C.

3. Distribution of quadrupole realization means

Previous work has shown the spatial correlation function has a minimum amount of sample variance (i.e., "cosmic variance") for a single frequency [33] or an equal-time (i.e., zero-lag) autocorrelation function [60,61]. According to this work, the (noise-free and infinite pulsar) angular power spectrum is a χ^2 distribution with $2(2\ell + 1)$ degrees of freedom for each multipole [33,60], with standard deviation $c_\ell / \sqrt{2\ell + 1}$, where c_ℓ is the value of the multipole coefficient from Eq. (10). In particular, the "cosmic variance" degrees of freedom for the quadrupole χ^2 distribution is 10, corresponding to a quadrupole standard deviation between realizations of approximately 0.14 for a single frequency.

In Fig. 9, we plot a histogram of the quadrupole posterior mean values for the three different realization studies. The orange curves in Fig. 9 show a χ^2 distribution with 10 degrees of freedom, which is the expected distribution from "cosmic variance" of the quadrupole. Each χ^2 distribution has been rescaled so that its mean value is equal to the injected value of the quadrupole given by Eq. (10). This comparison shows that our Bayesian harmonic analysis approach has a narrower distribution than expected from "cosmic variance", even with our modeling of noise and a finite number of pulsars.

A possible explanation for this difference is that our analysis uses all available data when determining the
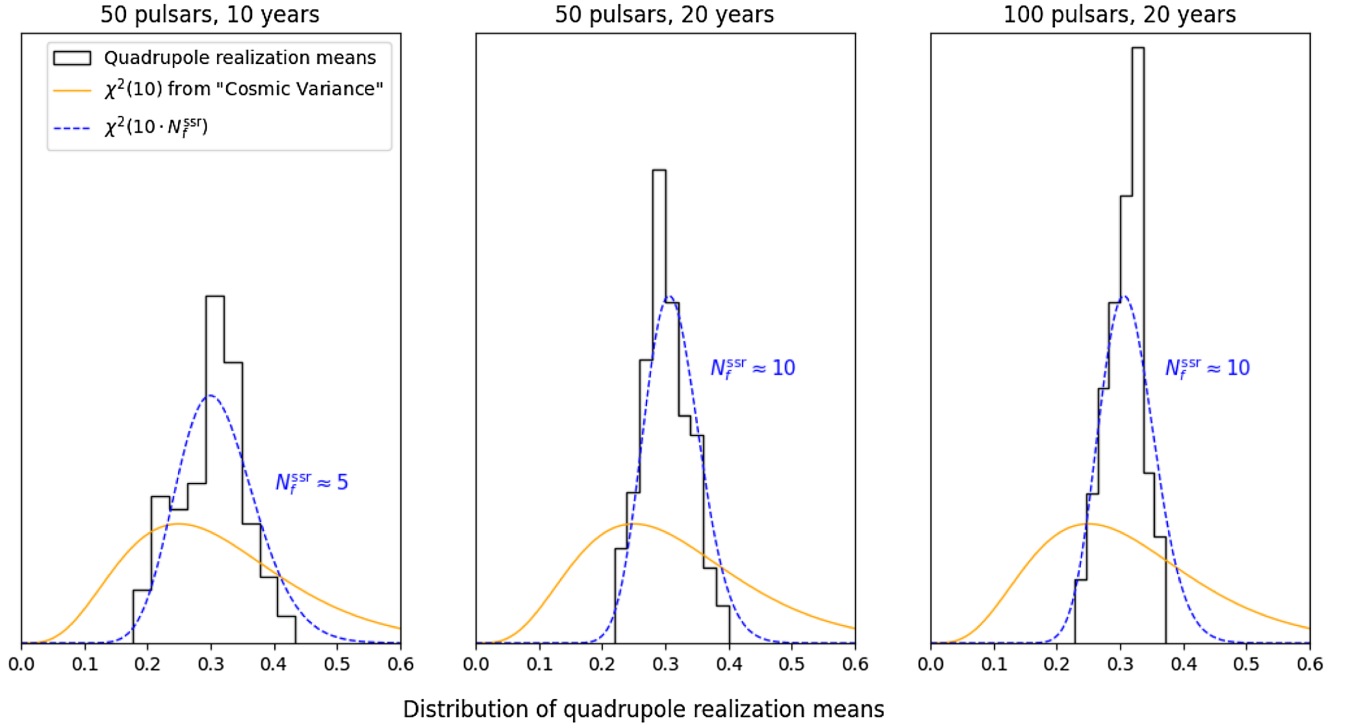
FIG. 9. Comparison of distribution of quadrupole realization-dependent means to a χ^2 distribution. Each of the three plots shows the results for a different realization study as indicated in the plot title. The less peaked χ^2 distribution (orange curves in each plot) shows the "cosmic variance" for a single frequency bin [33,60,61]. The "cosmic variance" degrees of freedom is 10 for the quadrupole χ^2 distribution, which we explain in the main text. The more sharply peaked χ^2 distribution (blue dashed curves) is obtained when the degrees of freedom scale linearly with the number of frequency bins in the strong-signal regime, N_f^{ssr}. The signal-dominated frequency bins are determined by counting the number of frequency bins in our mock data where the signal-to-noise ratio is at least 10 (using the average values of the injected source and noise signals). Each χ^2 distribution has been rescaled so that its mean value equals the value of the quadrupole coefficient from Eq. (10).

posterior distribution of the covariance matrix, as opposed to using a single frequency bin or an equal-time estimator. Since the sources that contribute to different frequencies of the SGWB are independent, each frequency bin provides an approximately independent sky map (see discussions in Sec. 4.3 of Ref. [33], Appendix C.3 of Ref. [60], and Appendix B.2 of Ref. [62]). Reference [33] suggests that the degrees of freedom of the multipole χ^2 distribution should scale with the number of frequency bins for a noise-free infinite pulsar dataset.

Since our analyses include intrinsic noise and use a finite number of pulsars, we expect only signal-dominated frequency bins to effectively contribute to these degrees of freedom. To determine the number of frequencies in the strong-signal regime, N_f^{ssr}, we count the number of frequency bins where the signal-to-noise ratio is at least 10 (using the average values of the injected source and noise signals). For our mock data, $N_f^{\mathrm{ssr}} \approx 5$ for 10 years of observation, and $N_f^{\mathrm{ssr}} \approx 10$ for 20 years of observation. Using a lower signal-to-noise ratio cutoff does not change these results appreciably.

The dashed blue curves in Fig. 9 show a χ^2 distribution with $10 \cdot N_f^{\mathrm{ssr}}$ degrees of freedom. We can see that our

realization study results are much more consistent with this distribution than with the orange, "cosmic variance", distribution. A similar result has been found in Ref. [63], where it was observed that the "cosmic variance" is consistent with a single-frequency model of the data, but combining multiple frequencies reduces the total variance (e.g., see Fig. 8 of Ref. [63]). We leave further investigations into this topic to future work.

APPENDIX B: SAVAGE-DICKEY MULTIPOLE EVIDENCE

We use the Savage-Dickey approach [50] to calculate a Bayes factor for each multipole in our harmonic analyses, which is a measure of the evidence for the multipole in our model. We calculate a Bayes factor by comparing the hypotheses $\mathcal{H}_{1,\ell}: c_\ell \neq 0$ against the hypothesis $\mathcal{H}_{2,\ell}: c_\ell = 0$. Our hypotheses are nested, $\mathcal{H}_{2,\ell} \subset \mathcal{H}_{1,\ell}$, because the prior for c_ℓ includes 0. Moreover, since the prior for c_ℓ is uniform on its range, the prior probability distribution is $p(c_\ell) = 1$ for all prior values of c_ℓ. The Savage-Dickey Bayes factor for multipole ℓ is therefore given by

$$\mathcal{B}_\ell = \frac{p(c_\ell = 0)}{p(c_\ell = 0 | d, \mathcal{H}_{1,\ell})} = \frac{1}{p(c_\ell = 0 | d, \mathcal{H}_{1,\ell})}, \quad \text{(B1)}$$

where d is the data and $p(c_\ell = 0 | d, \mathcal{H}_{1,\ell})$ is the marginalized 1D posterior distribution for parameter c_ℓ evaluated at $c_\ell = 0$.

APPENDIX C: CORNER PLOTS

In this appendix we present the corner plots for the nine harmonic analyses discussed in Sec. IV C. Figures 10–12 show the results of the harmonic analyses for 50, 100, and 150 pulsars, respectively. In each figure we overlay the contours for 10, 20, and 30 years of observation time.
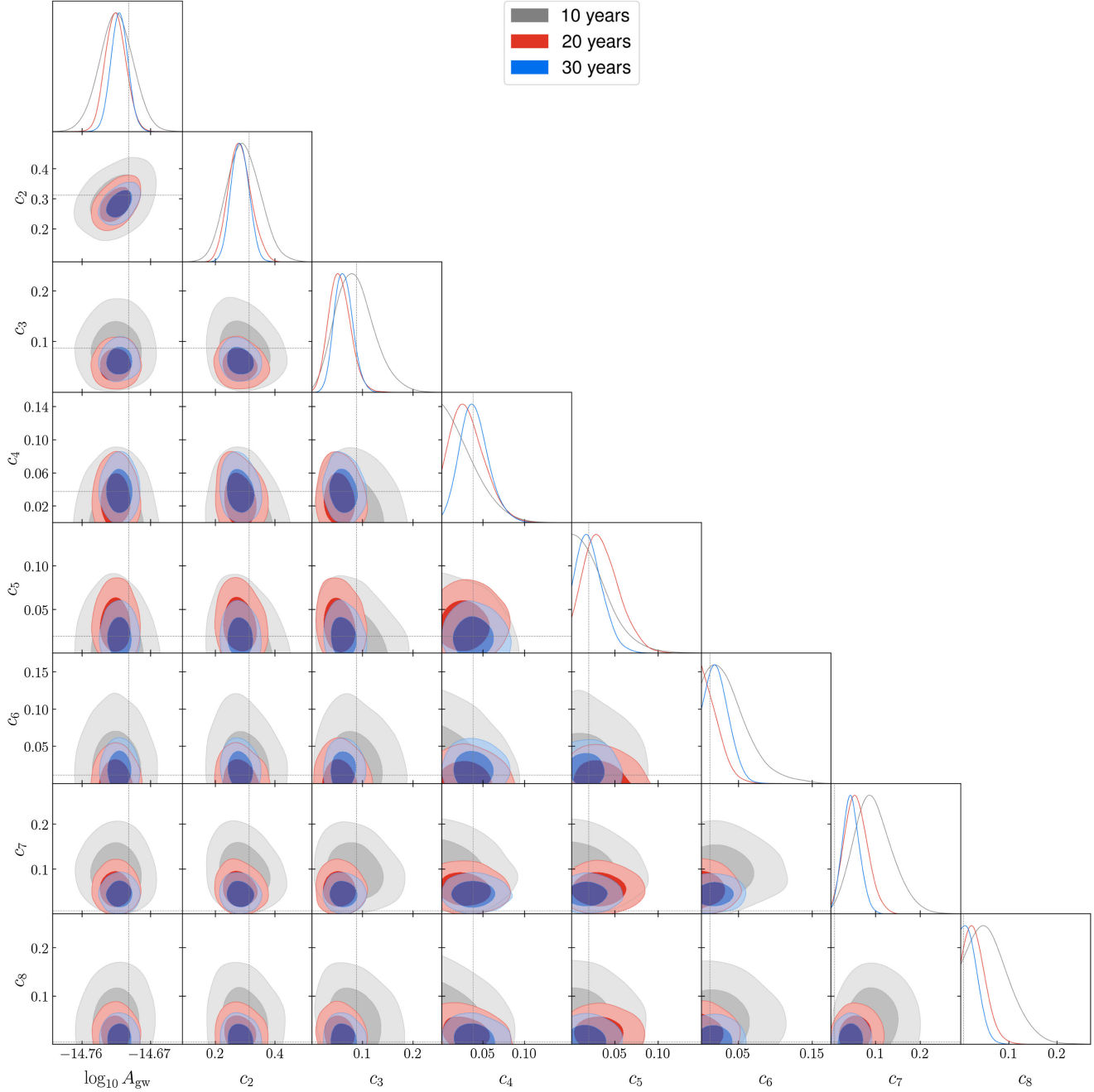

FIG. 10. Corner plots for harmonic analyses of mock PTA datasets involving 50 pulsars. The 68% and 95% CIs of the marginalized posteriors of $\log_{10} A_{\mathrm{gw}}$ and Legendre coefficients c_ℓ for total observation times of 10, 20, and 30 years are shown in gray, red, and blue, respectively. The injected values of the parameters are indicated by gray, dotted lines.
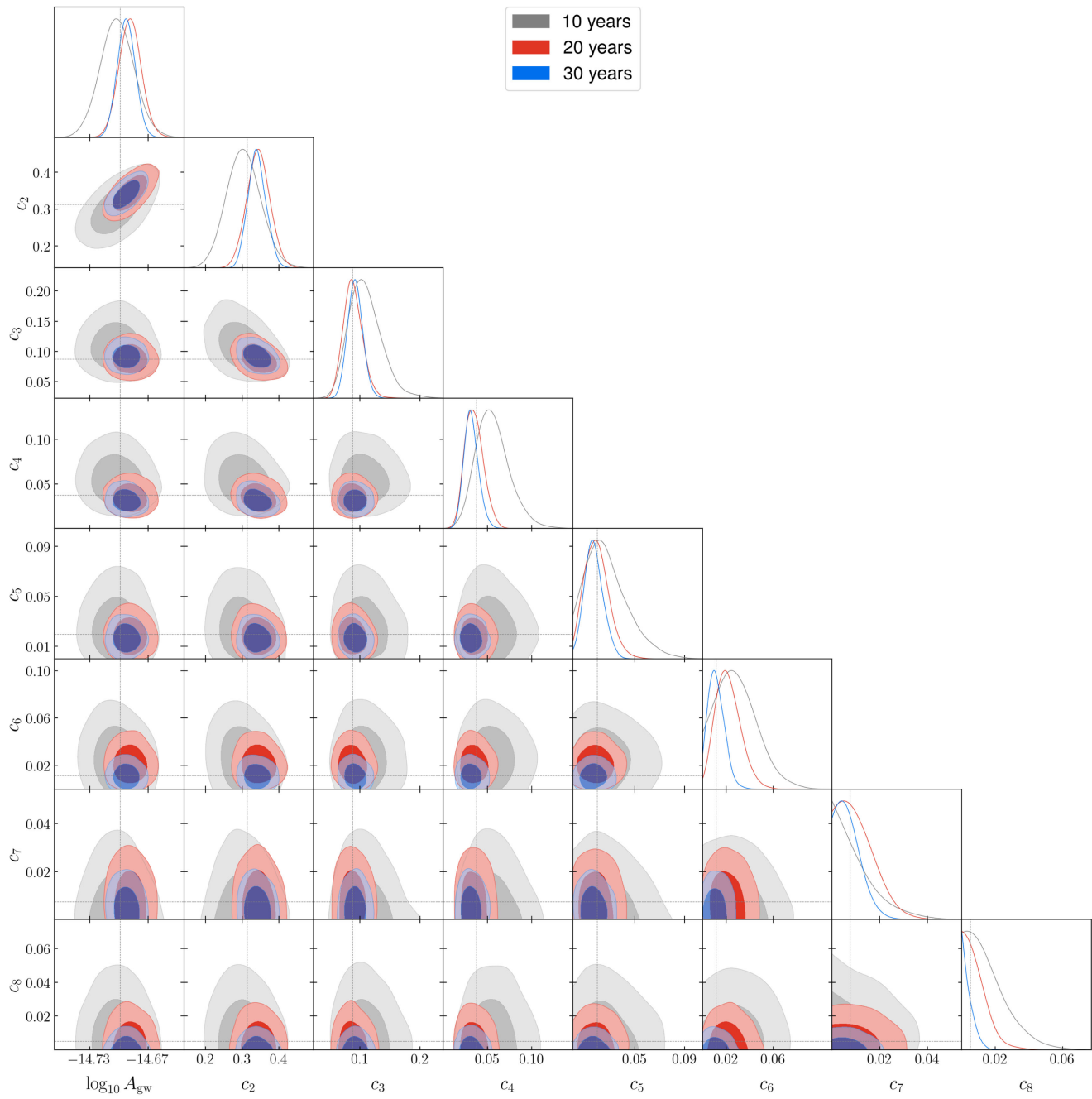
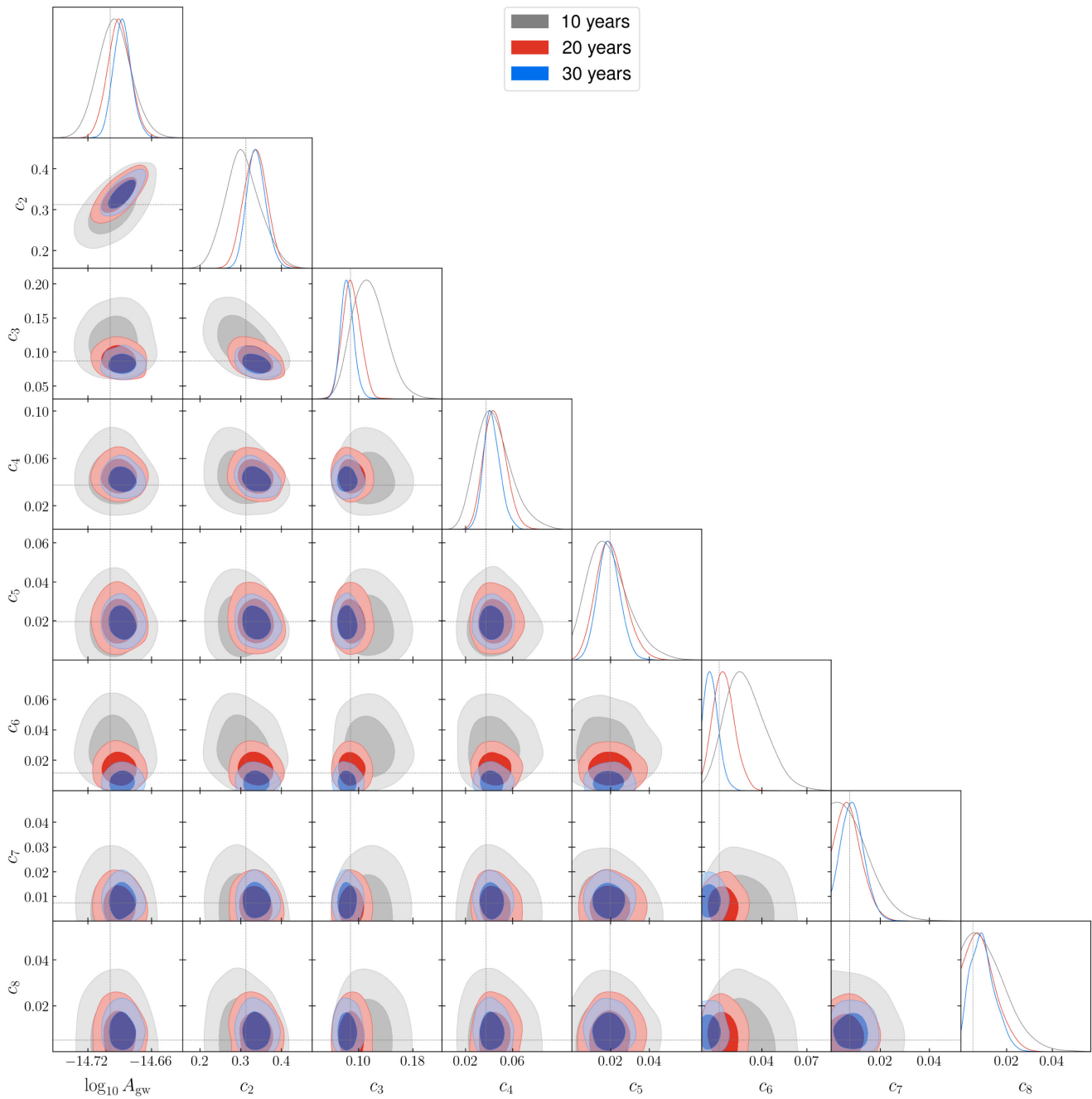FIG. 11. The same as Fig. 10, except with 100 pulsars.

FIG. 12. The same as Fig. 10, except with 150 pulsars.

[1] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Observation of gravitational waves from a binary black hole merger, Phys. Rev. Lett. **116**, 061102 (2016).

[2] K. K. Boddy *et al.*, Snowmass2021 theory frontier white paper: Astrophysical and cosmological probes of dark matter, J. High Energy Astrophys. **35**, 112 (2022).

[3] R. Caldwell *et al.*, Detection of early-universe gravitational-wave signatures and fundamental physics, Gen. Relativ. Gravit. **54**, 156 (2022).

[4] D. Green *et al.*, Snowmass theory frontier: Astrophysics and cosmology, arXiv:2209.06854.

[5] M. V. Sazhin, Opportunities for detecting ultralong gravitational waves, Sov. Astron. **22**, 36 (1978).

[6] S. L. Detweiler, Pulsar timing measurements and the search for gravitational waves, Astrophys. J. **234**, 1100 (1979).

[7] M. Maggiore, *Gravitational Waves. Vol. 2: Astrophysics and Cosmology* (Oxford University Press, New York, 2018).

[8] C. M. F. Mingarelli and J. A. Casey-Clyde, Seeing the gravitational wave universe, Science **378**, 592 (2022).

[9] R. W. Hellings and G. S. Downs, Upper limits on the isotropic gravitational radiation background from pulsar timing analysis, Astrophys. J. Lett. **265**, L39 (1983).

[10] J. Antoniadis *et al.*, The international pulsar timing array second data release: Search for an isotropic gravitational wave background, Mon. Not. R. Astron. Soc. **510**, 4873 (2022).

[11] Z. Arzoumanian *et al.* (NANOGrav Collaboration), The NANOGrav 12.5 yr data set: Search for an isotropic stochastic gravitational-wave background, Astrophys. J. Lett. **905**, L34 (2020).

[12] B. Goncharov *et al.*, On the evidence for a common-spectrum process in the search for the nanohertz gravitational-wave background with the Parkes pulsar timing array, Astrophys. J. Lett. **917**, L19 (2021).

[13] A. Chalumeau *et al.*, Noise analysis in the European pulsar timing array data release 2 and its implications on the gravitational-wave background search, Mon. Not. R. Astron. Soc. **509**, 5538 (2021).

[14] M. Anholm, S. Ballmer, J. D. E. Creighton, L. R. Price, and X. Siemens, Optimal strategies for gravitational wave stochastic background searches in pulsar timing data, Phys. Rev. D **79**, 084030 (2009).

[15] L. Dai, M. Kamionkowski, and D. Jeong, Total angular momentum waves for scalar, vector, and tensor fields, Phys. Rev. D **86**, 125013 (2012).

[16] W. Qin, K. K. Boddy, M. Kamionkowski, and L. Dai, Pulsar-timing arrays, astrometry, and gravitational waves, Phys. Rev. D **99**, 063002 (2019).

[17] J. Gair, J. D. Romano, S. Taylor, and C. M. F. Mingarelli, Mapping gravitational-wave backgrounds using methods from CMB analysis: Application to pulsar timing arrays, Phys. Rev. D **90**, 082001 (2014).

[18] D. P. Mihaylov, C. J. Moore, J. R. Gair, A. Lasenby, and G. Gilmore, Astrometric effects of gravitational wave backgrounds with non-Einsteinian polarizations, Phys. Rev. D **97**, 124058 (2018).

[19] D. P. Mihaylov, C. J. Moore, J. Gair, A. Lasenby, and G. Gilmore, Astrometric effects of gravitational wave backgrounds with nonluminal propagation speeds, Phys. Rev. D **101**, 024038 (2020).

[20] W. Qin, K. K. Boddy, and M. Kamionkowski, Subluminal stochastic gravitational waves in pulsar-timing arrays and astrometry, Phys. Rev. D **103**, 024045 (2021).

[21] R. C. Bernardo and K.-W. Ng, Stochastic gravitational wave background phenomenology in a pulsar timing array, Phys. Rev. D **107**, 044007 (2023).

[22] C. M. F. Mingarelli, T. Sidery, I. Mandel, and A. Vecchio, Characterizing gravitational wave stochastic background anisotropy with pulsar timing arrays, Phys. Rev. D **88**, 062005 (2013).

[23] S. R. Taylor and J. R. Gair, Searching for anisotropic gravitational-wave backgrounds using pulsar timing arrays, Phys. Rev. D **88**, 084001 (2013).

[24] S. C. Hotinli, M. Kamionkowski, and A. H. Jaffe, The search for anisotropy in the gravitational-wave background with pulsar-timing arrays, Open J. Astrophys. **2**, 8 (2019).

[25] E. S. Phinney, A practical theorem on gravitational wave backgrounds, arXiv:astro-ph/0108028.

[26] M. F. Alam *et al.* (NANOGrav Collaboration), The NANOGrav 12.5 yr data set: Wideband timing of 47 millisecond pulsars, Astrophys. J. Suppl. Ser. **252**, 5 (2021).

[27] M. F. Alam *et al.* (NANOGrav Collaboration), The NANOGrav 12.5 yr data set: Observations and narrowband timing of 47 millisecond pulsars, Astrophys. J. Suppl. Ser. **252**, 4 (2021).

[28] R. Blandford, R. Narayan, and R. W. Romani, Arrival-time analysis for a millisecond pulsar, J. Astrophys. Astron. **5**, 369388 (1984).

[29] L. Lentati, P. Alexander, M. P. Hobson, F. Feroz, R. van Haasteren, K. Lee, and R. M. Shannon, TempoNest: A Bayesian approach to pulsar timing analysis, Mon. Not. R. Astron. Soc. **437**, 3004 (2014).

[30] C. M. F. Mingarelli and T. Sidery, Effect of small interpulsar distances in stochastic gravitational wave background searches with pulsar timing arrays, Phys. Rev. D **90**, 062011 (2014).

[31] R. van Haasteren and M. Vallisneri, New advances in the Gaussian-process approach to pulsar-timing data analysis, Phys. Rev. D **90**, 104012 (2014).

[32] W. L. Burke, Large-scale random gravitational waves, Astrophys. J. **196**, 329 (1975).

[33] E. Roebber and G. Holder, Harmonic space analysis of pulsar timing array redshift maps, Astrophys. J. **835**, 21 (2017).

[34] Z. Arzoumanian *et al.* (NANOGrav Collaboration), The NANOGrav nine-year data set: Limits on the isotropic stochastic gravitational wave background, Astrophys. J. **821**, 13 (2016).

[35] N. S. Pol *et al.* (NANOGrav Collaboration), Astrophysics milestones for pulsar timing array gravitational-wave detection, Astrophys. J. Lett. **911**, L34 (2021).

[36] B. B. P. Perera *et al.*, The international pulsar timing array: Second data release, Mon. Not. R. Astron. Soc. **490**, 4666 (2019).

[37] G. Hobbs, R. Edwards, and R. Manchester, TEMPO2, a new pulsar timing package. 1. overview, Mon. Not. R. Astron. Soc. **369**, 655 (2006).

[38] https://github.com/vallis/libstempo.

[39] C. Xin, C. M. F. Mingarelli, and J. S. Hazboun, Multimessenger pulsar timing array constraints on supermassive black hole binaries traced by periodic light curves, Astrophys. J. **915**, 97 (2021).

[40] S. J. Chamberlin, J. D. E. Creighton, X. Siemens, P. Demorest, J. Ellis, L. R. Price, and J. D. Romano, Time-domain implementation of the optimal cross-correlation statistic for stochastic gravitational-wave background

searches in pulsar timing data, Phys. Rev. D **91**, 044048 (2015).

[41] J. A. Ellis, M. Vallisneri, S. R. Taylor, and P. T. Baker, Enterprise: Enhanced numerical toolbox enabling a robust pulsar inference suite (Zenodo, 2020), 10.5281/zenodo .4059815.

[42] S. R. Taylor, P. T. Baker, J. S. Hazboun, J. Simon, and S. J. Vigeland, ENTERPRISE-EXTENSIONS (2021), v2.3.3.

[43] J. Ellis and R. van Haasteren, jellis18/ptmcmcsampler: Official release (2017), 10.5281/zenodo.1037579.

[44] A. Gelman and D. B. Rubin, Inference from iterative simulation using multiple sequences, Stat. Sci. **7**, 457 (1992).

[45] J. S. Hazboun, J. Simon, X. Siemens, and J. D. Romano, Model dependence of Bayesian gravitational-wave background statistics for pulsar timing arrays, Astrophys. J. Lett. **905**, L6 (2020).

[46] E. Roebber, Ephemeris errors and the gravitational wave signal: Harmonic mode coupling in pulsar timing array searches, Astrophys. J. **876**, 55 (2019).

[47] R. van Haasteren, Y. Levin, P. McDonald, and T. Lu, On measuring the gravitational-wave background using pulsar timing arrays, Mon. Not. R. Astron. Soc. **395**, 1005 (2009).

[48] X. Siemens, J. Ellis, F. Jenet, and J. D. Romano, The stochastic background: Scaling laws and time to detection for pulsar timing arrays, Classical Quantum Gravity **30**, 224015 (2013).

[49] J. D. Romano, J. S. Hazboun, X. Siemens, and A. M. Archibald, Common-spectrum process versus cross-correlation for gravitational-wave searches using pulsar timing arrays, Phys. Rev. D **103**, 063027 (2021).

[50] J. M. Dickey, The weighted likelihood ratio, linear hypotheses on normal location parameters, Ann. Math. Stat. **42**, 204 (1971).

[51] B. P. Abbott et al. (LIGO Scientific, Virgo, Fermi-GBM, and INTEGRAL Collaborations), Gravitational waves and gamma-rays from a binary neutron star merger: GW170817 and GRB 170817A, Astrophys. J. Lett. **848**, L13 (2017).

[52] C. de Rham and S. Melville, Gravitational rainbows: LIGO and dark energy at its cutoff, Phys. Rev. Lett. **121**, 221101 (2018).

[53] Z. Arzoumanian et al. (NANOGrav Collaboration), The NANOGrav 12.5-year data set: Search for non-Einsteinian polarization modes in the gravitational-wave background, Astrophys. J. Lett. **923**, L22 (2021).

[54] C. Tiburzi, G. Hobbs, M. Kerr, W. Coles, S. Dai, R. Manchester, A. Possenti, R. Shannon, and X. You, A study of spatial correlations in pulsar timing array data, Mon. Not. R. Astron. Soc. **455**, 4339 (2016).

[55] G. Agazie et al. (NANOGrav Collaboration), The NANO-Grav 15 yr data set: Evidence for a gravitational-wave background, Astrophys. J. Lett. **951**, L8 (2023).

[56] J. Antoniadis et al. (EPTA and InPTA Collaborations), The second data release from the European pulsar timing array—III. Search for gravitational wave signals, Astron. Astrophys. **678**, A50 (2023).

[57] D. J. Reardon et al., Search for an isotropic gravitational-wave background with the parkes pulsar timing array, Astrophys. J. Lett. **951**, L6 (2023).

[58] H. Xu et al., Searching for the nano-hertz stochastic gravitational wave background with the Chinese pulsar timing array data release I, Res. Astron. Astrophys. **23**, 075024 (2023).

[59] A. Lewis, GetDist: A PYTHON package for analysing Monte Carlo samples, arXiv:1910.13970.

[60] B. Allen, Variance of the Hellings-Downs correlation, Phys. Rev. D **107**, 043018 (2023).

[61] B. Allen and J. D. Romano, Hellings and Downs correlation of an arbitrary set of pulsars, Phys. Rev. D **108**, 043026 (2023).

[62] Y. Ali-Haïmoud, T. L. Smith, and C. M. F. Mingarelli, Fisher formalism for anisotropic gravitational-wave background searches with pulsar timing arrays, Phys. Rev. D **102**, 122005 (2020).

[63] M. Caliskan, Y. Chen, L. Dai, N. Anil Kumar, I. Stomberg, and X. Xue, Dissecting stochastic gravitational wave background with astrometry, J. Cosmol. Astropart. Phys. 05 (2024) 030.