

Noise-marginalized optimal statistic: A robust hybrid frequentist-Bayesian statistic for the stochastic gravitational-wave background in pulsar timing arrays

Sarah J. Vigeland,¹ Kristina Islo,¹ Stephen R. Taylor,^{2,3} and Justin A. Ellis^{4,5,6}

¹*Center for Gravitation, Cosmology and Astrophysics, University of Wisconsin–Milwaukee, P.O. Box 413, Milwaukee, Wisconsin 53201, USA*

²*Theoretical Astrophysics Including Relativity (TAPIR), MC 350-17, California Institute of Technology, Pasadena, California 91125, USA*

³*Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California 91106, USA*

⁴*Department of Physics and Astronomy, West Virginia University, P.O. Box 6315, Morgantown, West Virginia 26506, USA*

⁵*Center for Gravitational Waves and Cosmology, West Virginia University, Chestnut Ridge Research Building, Morgantown, West Virginia 26505, USA*

⁶*Infinia ML, 202 Rigsbee Avenue, Durham, North Carolina 27701, USA*



(Received 1 June 2018; published 1 August 2018)

Observations have revealed that nearly all galaxies contain supermassive black holes (SMBHs) at their centers. When galaxies merge, these SMBHs form SMBH binaries (SMBHBs) that emit low-frequency gravitational waves (GWs). The incoherent superposition of these sources produce a stochastic GW background (GWB) that can be observed by pulsar timing arrays. The optimal statistic is a frequentist estimator of the amplitude of the GWB that specifically looks for the spatial correlations between pulsars induced by the GWB. In this paper, we introduce an improved method for computing the optimal statistic that marginalizes over the red noise in individual pulsars. We use simulations to demonstrate that this method more accurately determines the strength of the GWB, and we use the noise-marginalized optimal statistic to compare the significance of monopole, dipole, and Hellings-Downs (HD) spatial correlations and perform sky scrambles.

DOI: [10.1103/PhysRevD.98.044003](https://doi.org/10.1103/PhysRevD.98.044003)

I. INTRODUCTION

Long-wavelength gravitational waves (GWs) with frequencies of 10^{-9} – 10^{-7} Hz can be observed with pulsar timing arrays (PTAs) composed of millisecond pulsars (MSPs) [1,2]. The dominant astrophysical source in this frequency range is the isotropic stochastic gravitational wave background (GWB) made up of the incoherent superposition of GWs from inspiraling supermassive black hole binaries (SMBHBs) [3–5]. By monitoring the periodic emission from these pulsars using radio telescopes, we can probe the dynamics of the spacetime through which the pulses travel. This is done by searching for correlations in the pulsar timing residuals, which measure the differences between the expected and observed pulse times of arrival (TOAs). Current upper limits on the stochastic background from PTAs are approaching theoretical predictions for the GWB [6–8].

PTAs primarily use Bayesian data analysis to compare the inferred probabilities of various models for the residuals, including one where they contain the GWB [9,10]. Bayesian inference is a powerful tool because it properly

accounts for degeneracies between parameters and incorporates all sources of uncertainty into the analysis. However, running a full Bayesian analysis is computationally intensive, particularly when searching for evidence of Hellings-Downs (HD) spatial correlations—the “smoking gun” of the GWB.

The significance of the GWB can also be assessed using the optimal statistic, a frequentist estimator for the GWB amplitude [11–13]. Not only does it provide an independent detection procedure, complementing a more robust Bayesian analysis, but it requires significantly less time to compute. In particular, the optimal statistic produces results for a given spatial correlation function within seconds; a full Bayesian analysis including correlations has to run for many weeks on a supercomputing cluster.

However, when pulsars have significant red noise the optimal statistic gives biased results due to the strong covariance between the individual red noise parameters and the GWB amplitude. Many individual pulsars show evidence for red noise [14,15], and uncertainty in the position of the Solar System barycenter (SSB) leads to a common red process in all pulsars [8]. Here we present a technique

for improving the accuracy of the optimal statistic by including an additional step: marginalizing over the individual pulsars' red noise parameters using the posterior distributions from a full Bayesian analysis of all the pulsars. This hybrid approach produces a more precise estimate of the GWB amplitude A_{gw} and its uncertainty, while requiring only a few minutes more than the more traditional method of computation. Furthermore, the same Bayesian analysis drawn upon by the noise marginalization can be used to compute the optimal statistic for any choice of spatial correlations simply by changing the overlap-reduction function (ORF). For example, clock errors lead to a common red signal with monopole spatial correlations [16], while uncertainty in the SSB produces dipole spatial correlations [17]. This technique is used to perform the frequentist searches for common red signals with HD, monopole, and dipole spatial correlations in the NANOGrav 11-year data set [8].

This paper is organized as follows. In Sec. II we lay out the procedure for computing the noise-marginalized optimal statistic. We use simulations to compare the noise-marginalized optimal statistic to the optimal statistic computed with fixed noise. In Sec. III we determine how well the noise-marginalized optimal statistic can differentiate between monopole, dipole, and HD spatial correlations. In Sec. IV we use the noise-marginalized optimal statistic to perform sky scrambles, which assess the significance of HD spatial correlations by scrambling the pulsars' sky positions [18,19]. We summarize our results in Sec. V as well as discuss future applications of the noise-marginalized optimal statistic.

II. NOISE-MARGINALIZED OPTIMAL STATISTIC

The optimal statistic is a frequentist estimator for the amplitude of an isotropic stochastic GWB, and can be derived by analytically maximizing the PTA likelihood function in the weak-signal regime [11,13]. It is constructed from the timing residuals $\delta\mathbf{t}$, which can be written as

$$\delta\mathbf{t} = M\boldsymbol{\epsilon} + F\mathbf{a} + U\mathbf{j} + \mathbf{n}. \quad (1)$$

The term $M\boldsymbol{\epsilon}$ describes the contributions to the residuals from perturbations to the timing model. The term $U\mathbf{j}$ describes noise that is correlated for observations made at the same time at different frequencies and uncorrelated over different observing epochs, while \mathbf{n} describes uncorrelated white noise from TOA measurement uncertainties. The term $F\mathbf{a}$ describes red noise, including both red noise intrinsic to the pulsar and a common red noise signal common to all pulsars (such as a GW signal). We model the red noise as a Fourier series,

$$F\mathbf{a} = \sum_{j=1}^N \left[a_j \sin\left(\frac{2\pi jt}{T}\right) + b_j \cos\left(\frac{2\pi jt}{T}\right) \right], \quad (2)$$

where N is the number of Fourier modes used (typically $N = 30$) and T is the span of the observations.

The optimal statistic is constructed from the autocovariance and cross-covariance matrices C_a and S_{ab} ,

$$C_a = \langle \delta\mathbf{t}_a \delta\mathbf{t}_a^T \rangle, \quad (3)$$

$$S_{ab} = \langle \delta\mathbf{t}_a \delta\mathbf{t}_b^T \rangle|_{a \neq b}, \quad (4)$$

where $\delta\mathbf{t}_a$ is a vector of the residuals of the a th pulsar in the PTA. For the GWB with power spectral density (PSD) $P_{\text{gw}}(f)$ and overlap reduction function (ORF) Γ_{ab} , the cross-covariance matrices are

$$S_{ab} = F_a \phi_{ab}^{\text{gw}} F_b^T, \quad (5)$$

where

$$\phi_{ab}^{\text{gw}} = \Gamma_{ab} P_{\text{gw}}(f). \quad (6)$$

The ORF is the HD curve [1],

$$\Gamma_{ab} = \frac{1}{2} \left[1 - \frac{1}{2} \left(\frac{1 - \cos \theta_{ab}}{2} \right) + 3 \left(\frac{1 - \cos \theta_{ab}}{2} \right) \ln \left(\frac{1 - \cos \theta_{ab}}{2} \right) \right], \quad (7)$$

where θ_{ab} is the angle between the pulsars. We model the PSD of the GWB as a power law:

$$P_{\text{gw}}(f) = \frac{A_{\text{gw}}^2}{12\pi^2} \left(\frac{f}{f_{\text{yr}}} \right)^{-\gamma}, \quad (8)$$

where $\gamma = 13/3$ assuming SMBHBs evolve solely due to GW emission and $f_{\text{yr}} \equiv 1/(1 \text{ yr})$. The optimal statistic \hat{A}^2 is given by

$$\hat{A}^2 = \frac{\sum_{ab} \delta\mathbf{t}_a^T C_a^{-1} \tilde{S}_{ab} C_b^{-1} \delta\mathbf{t}_b}{\sum_{ab} \text{tr}(C_a^{-1} \tilde{S}_{ab} C_b^{-1} \tilde{S}_{ba})}, \quad (9)$$

where \tilde{S}_{ab} is the amplitude-independent cross-correlation matrix,

$$A_{\text{gw}}^2 \tilde{S}_{ab} = S_{ab}. \quad (10)$$

This definition of the optimal statistic ensures that $\langle \hat{A}^2 \rangle = A_{\text{gw}}^2$. If $A_{\text{gw}} = 0$, the variance of the optimal statistic is

$$\sigma_0 = \left[\sum_{ab} \text{tr}(C_a^{-1} \tilde{S}_{ab} C_b^{-1} \tilde{S}_{ba}) \right]^{-1/2}. \quad (11)$$

For a measured value of \hat{A}^2 , the significance of $\hat{A}^2 \neq 0$ is given by the signal-to-noise ratio (SNR)

TABLE I. Pulsar parameters used in simulated PTA data sets.

Pulsar	T_{obs} (yrs)	σ_w (μs)	A_{red}	γ_{red}
J0030 + 0451	11.0	0.339	-13.93	3.56
J0613 - 0200	11.0	0.281	-13.14	1.22
J1012 + 5307	11.0	0.320	-12.79	1.51
J1024 - 0719	6.0	0.421
J1455 - 3330	11.0	0.773
J1600 - 3053	8.0	0.146
J1614 - 2230	7.0	0.261
J1640 + 2224	11.0	0.202
J1713 + 0747	11.0	0.093	-14.14	1.58
J1741 + 1351	6.0	0.106
J1744 - 1134	11.0	0.096
B1855 + 09	11.0	0.218	-13.75	3.54
J1853 + 1303	7.0	0.215
J1909 - 3744	11.0	0.034	-13.84	1.74
J1918 - 0642	11.0	0.342
J2010 - 1323	7.0	0.413
J2145 - 0750	11.0	0.281	-12.69	1.30
J2317 + 1439	11.0	0.160

$$\rho = \frac{\sum_{ab} \delta \mathbf{t}_a^T C_a^{-1} \tilde{S}_{ab} C_b^{-1} \delta \mathbf{t}_b}{[\sum_{ab} \text{tr}(C_a^{-1} \tilde{S}_{ab} C_b^{-1} \tilde{S}_{ba})]^{1/2}}. \quad (12)$$

When constructing the residuals $\delta \mathbf{t}_a$, we typically fix the red noise parameters to the values that maximize the single-pulsar likelihood. However, this leads to a bias in the optimal statistic because the individual red noise and common red noise parameters are highly covariant, with the optimal statistic computed using fixed red noise parameters systematically lower than the true value of A_{gw}^2 . In this section, we compare three techniques for computing the optimal statistic. First, we fix the individual pulsars' red noise parameters to the maximum-likelihood values from individual Bayesian pulsar noise analyses. Second, we fix the pulsars' red noise parameters to the values that jointly maximize the likelihood for a Bayesian analysis of all of the pulsars in our PTA that searches over the pulsars' red noise parameters and a common red process. For the noise-marginalized method, we draw values of the pulsars' red noise parameters from the posteriors generated by the common Bayesian analysis.

We use these methods to compute the optimal statistic for simulated ‘‘NANOGrav-like’’ data sets consisting of 18 MSPs with observation times, sky positions, and noise properties matching the 18 longest-observed pulsars in the NANOGrav 11-year data set [15]. We include white noise for all pulsars, plus red noise parametrized as a power law,

$$P_a(f) = \frac{A_{\text{red}}^2}{12\pi^2} \left(\frac{f}{f_{\text{yr}}} \right)^{-\gamma}, \quad (13)$$

for those pulsars that show evidence of red noise (see Table I for more details). We use the PTA data analysis package PAL2 [20] to perform the noise analyses and compute the optimal statistic.

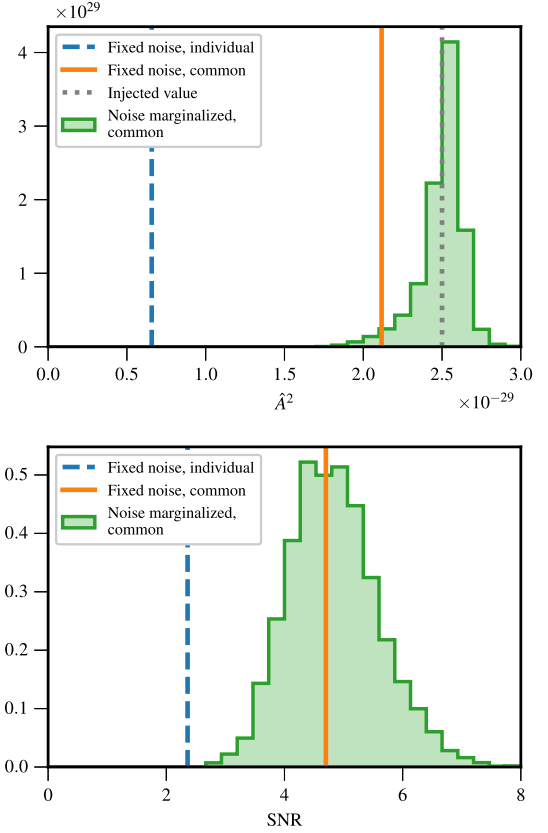


FIG. 1. Optimal statistic for a simulated PTA data set containing a GWB with $A_{\text{gw}} = 5 \times 10^{-15}$. The fixed-noise analysis using the individual noise values (dashed blue lines) systematically underestimates \hat{A}^2 , while the fixed-noise analysis using the common noise values (solid orange lines) and the noise-marginalized analysis (green histograms) more accurately recover A_{gw} .

Figure 1 shows the fixed-noise and noise-marginalized optimal statistic for a simulation with a GWB with $A_{\text{gw}} = 5 \times 10^{-15}$. For this particular realization of the GWB, the fixed-noise analysis using the individual noise results gives $\hat{A}^2 = 6.6 \times 10^{-30}$ with $\text{SNR} = 2.4$, and the fixed-noise analysis using the common noise results gives $\hat{A}^2 = 2.1 \times 10^{-29}$ with $\text{SNR} = 4.7$. The noise-marginalized analysis gives $\hat{A}^2 = (2.5 \pm 0.1) \times 10^{-29}$ with $\text{SNR} = 4.8 \pm 0.8$. The value of \hat{A}^2 from the fixed-noise analysis using the individual noise results is significantly lower than the injected level of the GWB, while the values of \hat{A}^2 from the fixed-noise analysis using the common noise results and the noise-marginalized analysis are in good agreement with each other and the injected value. The fixed-noise analysis using the individual noise results also gives a significantly lower SNR than the other two.

In Fig. 2 we show the optimal statistic for 300 different realizations of a GWB with $A_{\text{gw}} = 5 \times 10^{-15}$ computed using the three techniques described above. For the noise-marginalized analysis, we plot the mean values of \hat{A}^2 and ρ .

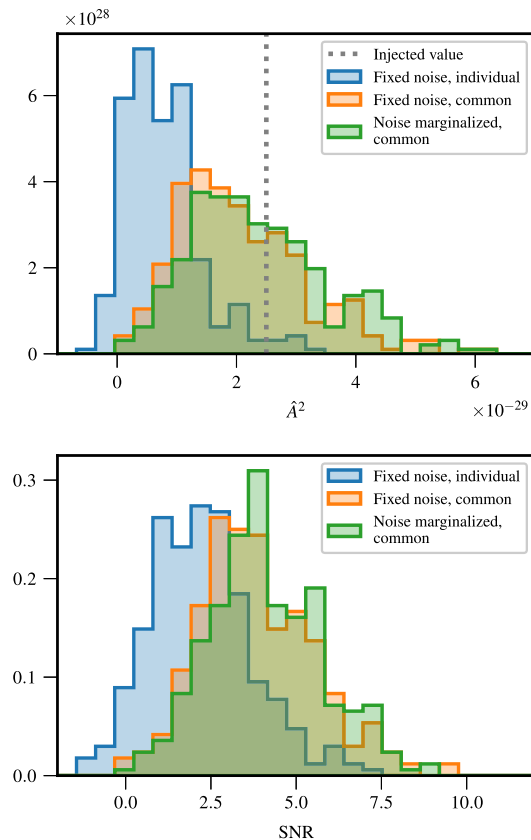


FIG. 2. Optimal statistic and SNR for 300 simulated data sets containing a GWB with $A_{\text{gw}} = 5 \times 10^{-15}$. The fixed-noise analysis using the individual noise values (blue) systematically underestimates \hat{A}^2 , while both the fixed-noise analysis using the common noise values (orange) and the noise-marginalized analysis (green) accurately recover A_{gw} .

Using the noise values from individual noise analyses systematically underestimates the strength of the GWB, while using the noise values from a common noise analysis more accurately recovers the injected value. The fixed-noise analysis using the individual noise results finds $\hat{A}^2 = (7.9 \pm 6.8) \times 10^{-30}$ and $\rho = 2.3 \pm 1.5$, averaging over realizations of the GWB. The fixed-noise and noise-marginalized analyses using the common noise results both give $\hat{A}^2 = (2.4 \pm 1.2) \times 10^{-29}$ and $\rho = 4.1 \pm 1.7$.

The fixed-noise and noise-marginalized analyses using the common noise results give the same results for $A_{\text{gw}} = 5 \times 10^{-15}$, but there are differences between them when analyzing data sets containing smaller injected values of A_{gw} . In Fig. 3 we show a P-P plot of the cumulative fraction of simulations for which the injected A_{gw}^2 lies within a given confidence interval of the measured \hat{A}^2 . The confidence interval of \hat{A}^2 is determined assuming \hat{A}^2 follows a Gaussian distribution, with mean and variance $\sigma_{\hat{A}^2}^2$ taken from the distribution for \hat{A}^2 found from our 300 realizations of the GWB (i.e., the top panel of Fig. 2). If \hat{A}^2 has a Gaussian distribution centered around A_{gw}^2 , the curves should lie along a straight line with slope equal to unity (the dotted, diagonal lines in Fig. 3).

We compare the three methods for computing the optimal statistic for simulations with $A_{\text{gw}} = 5 \times 10^{-15}$, $A_{\text{gw}} = 10^{-15}$, and $A_{\text{gw}} = 5 \times 10^{-16}$. The fixed-noise optimal statistic using the individual noise results systematically underestimates \hat{A}^2 (Fig. 3, left panel). The fixed-noise optimal statistic using the common noise results recovers \hat{A}^2 well for large values of A_{gw} , but for small values it also underestimates \hat{A}^2 (Fig. 3,

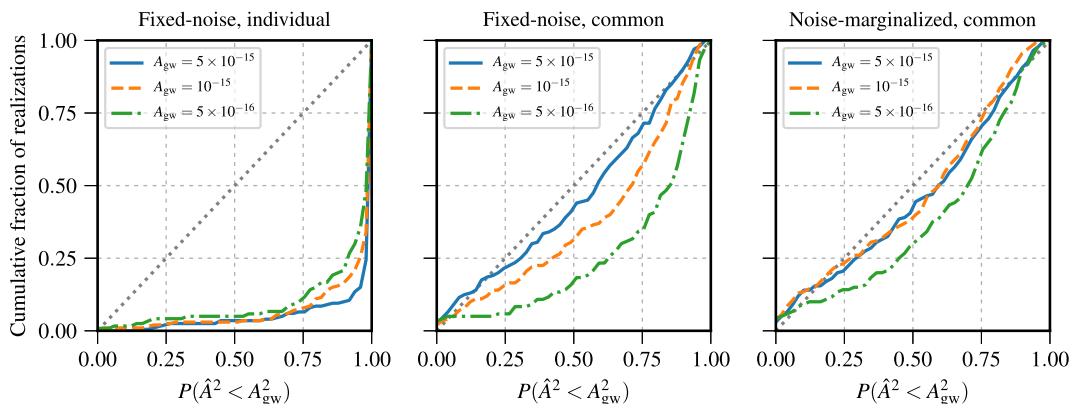


FIG. 3. P-P plot showing the cumulative fraction of simulations for which A_{gw}^2 lies within a given confidence interval of the measured \hat{A}^2 . The probability distribution of \hat{A}^2 is assumed to be a Gaussian with variance $\sigma_{\hat{A}^2}^2$. The fixed-noise optimal statistic using the individual and common noise results both give biased values of \hat{A}^2 , particularly for small values of A_{gw} , while the noise-marginalized optimal statistic gives more accurate values of \hat{A}^2 over a large range of injected values of A_{gw} .

middle panel). The noise-marginalized optimal statistic provides the most accurate estimate of \hat{A}^2 over the range of A_{gw} considered here, although it still slightly underestimates A_{gw} (Fig. 3, right panel).

III. MONOPOLE AND DIPOLE SPATIAL CORRELATIONS

The optimal statistic is particularly well suited to compare multiple spatial correlation relations because using a different spatial correlation only requires changing the ORF in Eq. (6). Tiburzi *et al.* [21] demonstrated how the optimal statistic can be altered to fit for multiple spatial correlations at once in order to mitigate common noise sources such as clock error and ephemeris error. Here we take a different approach—rather than simultaneously fitting for signals with different spatial correlations, we look at how well we can distinguish between different spatial correlations by computing the optimal statistic with monopole and dipole spatial correlations for the same simulated data sets as in the previous section. For a monopole signal, the ORF becomes

simply $\Gamma_{ab} = 1$, while for a dipole signal, the ORF becomes $\Gamma_{ab} = \cos \theta_{ab}$.

Our ability to distinguish between different spatial correlations depends on the strength of the GWB and the angular separations between pulsar pairs, θ_{ab} . We can determine the overlap between ORFs corresponding to different spatial correlations by computing the “match statistic” [18],

$$\bar{M} = \frac{\sum_{a,b \neq a} \Gamma_{ab} \Gamma'_{ab}}{\sqrt{(\sum_{a,b \neq a} \Gamma_{ab} \Gamma_{ab})(\sum_{a,b \neq a} \Gamma'_{ab} \Gamma'_{ab})}}, \quad (14)$$

where Γ and Γ' are two different ORFs. For the 18 pulsars used in these simulations, the match statistic for monopole and HD correlations is $\bar{M} = 0.264$, and the match statistic for dipole and HD correlations is $\bar{M} = 0.337$. These match statistics describe a fundamental limit on our ability to identify the spatial correlations of a common red signal as HD rather than monopole or dipole that depends only on the number of pulsars in our PTA and their sky positions.

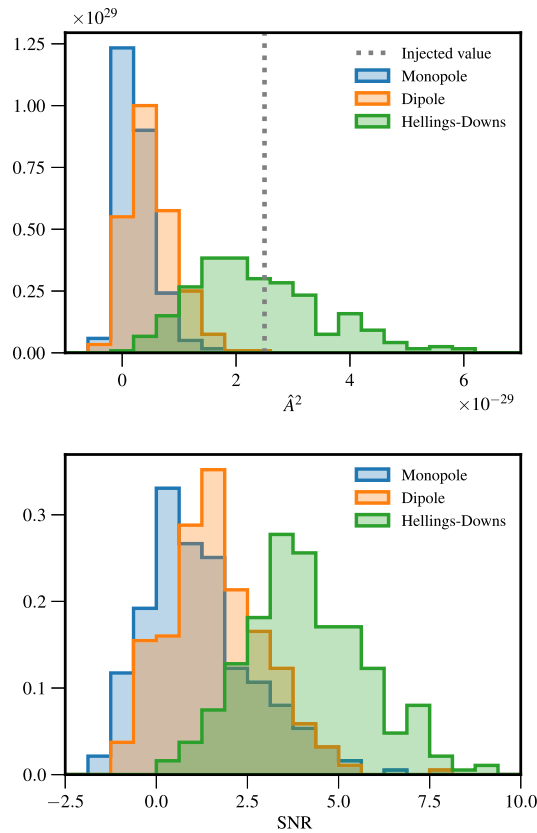


FIG. 4. Noise-marginalized mean optimal statistic and mean SNR for 300 simulated data sets containing an injected stochastic signal with dipole spatial correlations and $A = 5 \times 10^{-15}$. We compare the values of the mean optimal statistic and the SNR found using monopole (blue), dipole (orange), and HD (green) spatial correlations. The dashed vertical line indicates the injected value, $\hat{A}^2 = 2.5 \times 10^{-29}$.

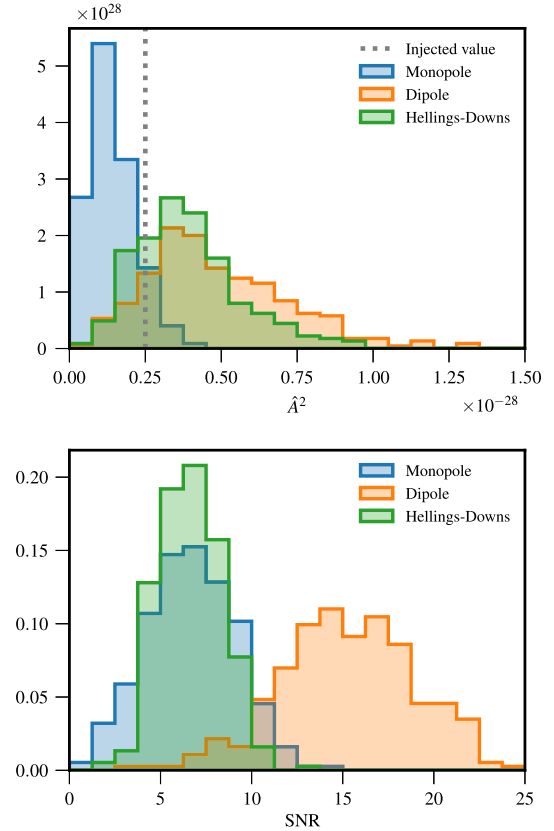


FIG. 5. Noise-marginalized mean optimal statistic and mean SNR for 300 simulated data sets containing an injected stochastic signal with dipole spatial correlations and $A = 5 \times 10^{-15}$. We compare the values of the optimal statistic and mean SNR found using monopole (blue), dipole (orange), and HD (green) spatial correlations. The dashed vertical line indicates the injected value, $\hat{A}^2 = 2.5 \times 10^{-29}$.

Figure 4 shows the noise-marginalized mean value of \hat{A}^2 and the mean SNR computed assuming monopole, dipole, and HD spatial correlations for 300 simulated data sets. Using a monopole or dipole ORF gives a lower value for the mean optimal statistic and mean SNR compared to the HD ORF. Using HD spatial correlations gives $\hat{A}^2 = (2.4 \pm 1.1) \times 10^{-29}$, while using monopole spatial correlations gives $\hat{A}^2 = (2.5 \pm 3.2) \times 10^{-30}$, and dipole spatial correlations gives $\hat{A}^2 = (5.2 \pm 4.4) \times 10^{-30}$. We find a noise-marginalized mean SNR above 1.0 in 97% of our simulated data sets using the HD ORF, and in 50% and 68% of our simulated data sets using the monopole and dipole ORFs, respectively. The mean SNR using the HD ORF, averaged over realizations of the GWB, is 4.1, and we find an SNR greater than this using the monopole and dipole ORFs in just 3% and 3.5% of our simulations, respectively.

This overlap between the monopole, dipole, and HD ORFs also means that a common red process that does not have HD correlations may be confused for a GWB. Figure 5 shows the results of 300 simulations containing a stochastic signal with dipole spatial correlations. Although a dipole signal has been injected, the HD ORF gives a mean SNR greater than 5 in 82% of the simulations. However, both the monopole and HD ORFs give smaller values of the mean \hat{A}^2 and mean SNR compared to the dipole ORF. Furthermore, there are no simulations for which the mean SNR with HD ORF is greater than the mean SNR with dipole ORF. This demonstrates the importance of comparing the SNR from different spatial correlations when determining the type of spatial correlations present.

IV. SKY SCRAMBLES

The significance of spatial correlations can also be tested with “sky scrambles,” where the ORF is altered in order to simulate changing the pulsars’ positions [18,19]. The scrambled ORFs are required to have small values of \bar{M} with the true ORF and each other so that they form a nearly orthogonal set. This ensures that the distribution of \hat{A}^2 computed using the scrambled ORFs forms the null hypothesis distribution. Taylor *et al.* [19] showed how sky scrambles affect the Bayes’ factor for simulated data sets. We performed a similar analysis using frequentist methods.

We generated 725 scrambled ORFs using a Monte Carlo algorithm. We required the scrambled ORFs to have $\bar{M} < 0.2$ with respect to the true ORF and each other. This threshold was chosen to be comparable to the match statistics between the HD ORF with monopole and dipole ORFs given in Sec. III. We did not choose a smaller threshold because significantly more time would have been needed to generate 725 scrambled ORFs. For each simulation, we computed the noise-marginalized mean optimal statistic and mean SNR for each scrambled ORF, and compared them to the values found using the true ORF.

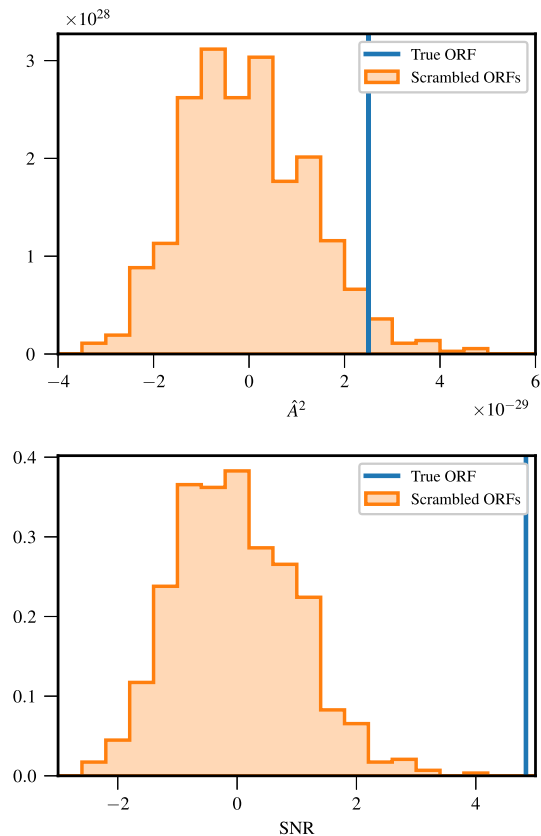


FIG. 6. Comparison between the noise-marginalized mean optimal statistic and mean SNR with and without sky scrambles for a simulated data set containing a GWB with $A_{\text{gw}} = 5 \times 10^{-15}$. None of the 725 scrambled ORFs gave a mean SNR as large as the mean SNR using the true ORF ($p < 0.0014$).

Figure 6 shows the results of a sky scramble analysis for a sample data set with $A_{\text{gw}} = 5 \times 10^{-15}$. For this particular realization of the GWB, none of the 725 scrambled ORFs resulted in a mean SNR greater than the mean SNR using

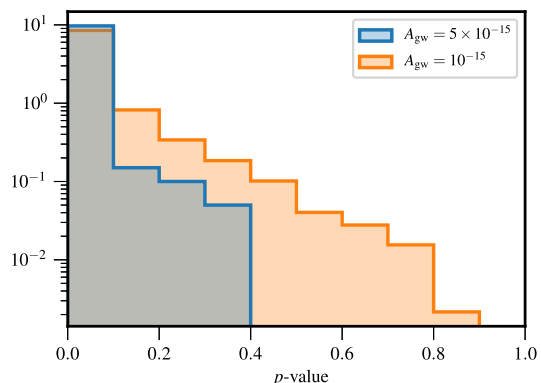


FIG. 7. Distribution of p -values for the noise-marginalized optimal statistic mean SNR using the true ORF compared to 725 sky scrambles from 300 realizations of the GWB. We show results for simulations with $A_{\text{gw}} = 5 \times 10^{-15}$ (blue) and $A_{\text{gw}} = 10^{-15}$ (orange).

the true ORF ($p < 0.0014$). In Fig. 7, we plot the distribution of p -values of the 725 sky scrambles for 300 realizations of the GWB. For a GWB with $A_{\text{gw}} = 5 \times 10^{-15}$, 95% of the simulations have $p \leq 0.05$, and 74% of the simulations have $p \leq 0.003$. For a GWB with $A_{\text{gw}} = 10^{-15}$, 76% of the simulations have $p \leq 0.05$, and 39% have $p \leq 0.003$. This shows that for smaller values of A_{gw} , there is a greater chance that noise fluctuations will appear to have the spatial correlations of the GWB.

V. CONCLUSION

The definitive signature of a GWB in PTA data is spatial correlations described by the HD curve. Searching for these using a full Bayesian approach is computationally expensive, requiring many weeks on a supercomputing cluster. In contrast, the optimal statistic can be computed in seconds. In this paper, we introduce an improved method for computing the optimal statistic, which uses the output from a Bayesian analysis for individual and common red signals to marginalize the optimal statistic over the individual pulsars' red noises. As shown in Sec. II, the noise-marginalized optimal statistic more accurately recovers the GWB amplitude than the fixed-noise optimal statistic, which underestimates the GWB amplitude when significant red noise is present in some pulsars.

Although the noise-marginalized optimal statistic requires computing the optimal statistic thousands of times, it is still many orders of magnitude faster than a Bayesian search. Furthermore, the results from a single Bayesian analysis, which are needed to marginalize over the red noise parameters, can be used to compute the optimal statistic for many different spatial correlations. In Sec. III we use the noise-marginalized optimal statistic to compare

the strength of monopole, dipole, and HD correlations in simulated PTA data with a GWB. In Sec. IV we use the noise-marginalized optimal statistic to perform sky scramble analyses, where we compare the mean SNR computed using the true ORF to the mean SNR computed using scrambled ORFs and measure the significance of HD spatial correlations through the p -value.

The primary strength of the optimal statistic is how quickly it can be computed. This is useful for analyses where the significance of many spatial correlations is compared, as with the sky scrambles. An upcoming paper will use the noise-marginalized optimal statistic to determine how well the spatial correlations corresponding to alternate GW polarizations can be measured. It also makes the optimal statistic a valuable tool for analyzing simulations where many realizations of the GWB are compared. The noise marginalization technique described in this paper is key to being able to accurately measure the GWB with the optimal statistic for real PTAs and realistic PTA simulations, for which red noise is significant.

ACKNOWLEDGMENTS

We thank Joe Lazio, Andrea Lommen, Joe Romano, Xavier Siemens, and Jolien Creighton for useful discussions. This work was supported by NSF Physics Frontier Center Grant No. 1430284. J. A. E. was partially supported by NASA through Einstein Fellowship Grant No. PF4-150120. We are grateful for computational resources provided by the Leonard E. Parker Center for Gravitation, Cosmology and Astrophysics at the University of Wisconsin–Milwaukee, which is supported by NSF Grants No. 0923409 and No. 1626190.

-
- [1] R. W. Hellings and G. S. Downs, *Astrophys. J. Lett.* **265**, L39 (1983).
 - [2] R. S. Foster and D. C. Backer, *Astrophys. J.* **361**, 300 (1990).
 - [3] M. Rajagopal and R. W. Romani, *Astrophys. J.* **446**, 543 (1995).
 - [4] A. H. Jaffe and D. C. Backer, *Astrophys. J.* **583**, 616 (2003).
 - [5] J. S. B. Wyithe and A. Loeb, *Astrophys. J.* **590**, 691 (2003).
 - [6] R. M. Shannon, V. Ravi, W. A. Coles, G. Hobbs, M. J. Keith, R. N. Manchester, J. S. B. Wyithe, M. Bailes, N. D. R. Bhat, S. Burke-Spolaor *et al.*, *Science* **342**, 334 (2013).
 - [7] L. Lentati, S. R. Taylor, C. M. F. Mingarelli, A. Sesana, S. A. Sanidas, A. Vecchio, R. N. Caballero, K. J. Lee, R. van Haasteren, S. Babak *et al.*, *Mon. Not. R. Astron. Soc.* **453**, 2576 (2015).
 - [8] Z. Arzoumanian, P. T. Baker, A. Brazier, S. Burke-Spolaor, S. J. Chamberlin, S. Chatterjee, B. Christy, J. M. Cordes, N. J. Cornish, F. Crawford *et al.*, *Astrophys. J.* **859**, 47 (2018).
 - [9] R. van Haasteren, Y. Levin, P. McDonald, and T. Lu, *Mon. Not. R. Astron. Soc.* **395**, 1005 (2009).
 - [10] L. Lentati, P. Alexander, M. P. Hobson, S. Taylor, J. Gair, S. T. Balan, and R. van Haasteren, *Phys. Rev. D* **87**, 104021 (2013).
 - [11] M. Anholm, S. Ballmer, J. D. E. Creighton, L. R. Price, and X. Siemens, *Phys. Rev. D* **79**, 084030 (2009).
 - [12] P. B. Demorest, R. D. Ferdman, M. E. Gonzalez, D. Nice, S. Ransom, I. H. Stairs, Z. Arzoumanian, A. Brazier, S. Burke-Spolaor, S. J. Chamberlin *et al.*, *Astrophys. J.* **762**, 94 (2013).
 - [13] S. J. Chamberlin, J. D. E. Creighton, X. Siemens, P. Demorest, J. Ellis, L. R. Price, and J. D. Romano, *Phys. Rev. D* **91**, 044048 (2015).

- [14] M. T. Lam, J. M. Cordes, S. Chatterjee, Z. Arzoumanian, K. Crowter, P. B. Demorest, T. Dolch, J. A. Ellis, R. D. Ferdman, E. Fonseca *et al.*, *Astrophys. J.* **834**, 35 (2017).
- [15] Z. Arzoumanian, A. Brazier, S. Burke-Spolaor, S. Chamberlin, S. Chatterjee, B. Christy, J. M. Cordes, N. J. Cornish, F. Crawford, H. T. Cromartie *et al.*, *Astrophys. J. Suppl. Ser.* **235**, 37 (2018).
- [16] G. Hobbs, W. Coles, R. N. Manchester, M. J. Keith, R. M. Shannon, D. Chen, M. Bailes, N. D. R. Bhat, S. Burke-Spolaor, D. Champion *et al.*, *Mon. Not. R. Astron. Soc.* **427**, 2780 (2012).
- [17] D. J. Champion, G. B. Hobbs, R. N. Manchester, R. T. Edwards, D. C. Backer, M. Bailes, N. D. R. Bhat, S. Burke-Spolaor, W. Coles, P. B. Demorest *et al.*, *Astrophys. J. Lett.* **720**, L201 (2010).
- [18] N. J. Cornish and L. Sampson, *Phys. Rev. D* **93**, 104047 (2016).
- [19] S. R. Taylor, L. Lentati, S. Babak, P. Brem, J. R. Gair, A. Sesana, and A. Vecchio, *Phys. Rev. D* **95**, 042002 (2017).
- [20] J. Ellis and R. van Haasteren, jellis18/pal2: Pal2, 2017, <https://doi.org/10.5281/zenodo.251456>.
- [21] C. Tiburzi, G. Hobbs, M. Kerr, W. A. Coles, S. Dai, R. N. Manchester, A. Possenti, R. M. Shannon, and X. P. You, *Mon. Not. R. Astron. Soc.* **455**, 4339 (2016).