

Optimal strategies for continuous wave detection in pulsar timing arrays: Realistic pulsar noise and a gravitational wave background

Shashwat C. Sardesai^{1,2}, Gabriel E. Freedman¹, Sarah J. Vigeland¹ and Caitlin A. Witt³

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

²*Department of Physics and Astronomy, Texas Tech University, Lubbock, Texas 79409, USA*

³*Department of Physics, Wake Forest University, Winston-Salem, North Carolina 27103, USA*



(Received 1 November 2025; accepted 7 April 2026; published 30 April 2026)

Pulsar timing arrays are sensitive to low-frequency gravitational waves (GWs), such as those produced by supermassive binary black holes at subparsec separations. The incoherent superposition of GWs emitted by a cosmological population of these sources produces a gravitational wave background (GWB), while some individual sources may be resolvable as deterministic signals with slowly varying GW frequencies, which are often referred to as “continuous waves” (CWs). The $\mathcal{F}_p$ -statistic is a frequentist method of detecting these CWs. In this paper, we study how the presence of pulsar red noise and a GWB affect the $\mathcal{F}_p$ -statistic. We compare results when marginalizing over the red noise and using the maximum-likelihood values of the red noise, and find little difference between the two. We also present results of using the $\mathcal{F}_p$ -statistic to analyze the NANOGrav 12.5-year dataset, where we find no evidence for CWs in agreement with the previously published Bayesian results.

DOI: 10.1103/174h-7xf1

I. INTRODUCTION

Pulsar timing arrays (PTAs) use millisecond pulsars to detect the presence of low-frequency gravitational waves (GWs) [1–3]. When a passing GW passes between the line-of-sight between the Earth and the pulsar, it induces a redshift which results in a delay in the times of arrival of the pulsar signals. By measuring the time-delay (or residuals) δt , we can detect the presence of GWs. Millisecond pulsars are ideal for this purpose owing to their exceptional rotational stability—they can be timed to submicrosecond precision over decades [4,5]. Many astrophysical and cosmological sources are predicted to generate low-frequency GWs that can be observed by PTAs, including supermassive black hole binaries (SMBHBs) at subparsec separations ($f \sim 10^{-9}$ – 10^{-7} Hz) [6–9]. Recently, PTA Collaborations around the world have published the first evidence of nanohertz GWs, which is consistent with predictions for a gravitational wave background (GWB) produced by a population of SMBHBs [10–14].

If the GWB is coming from a population of SMBHBs, we expect that some binaries can be resolved as individual sources [15]. These sources would appear as

continuous waves (CWs) because the GW frequency evolves very slowly compared to the timespan of the dataset. So far searches of NANOGrav datasets have found no evidence of CWs [16–18], but we expect that future PTA datasets will detect an individual binary within the next 5–10 years [15,19].

PTAs have primarily performed Bayesian searches for CWs, in which the pulsars’ intrinsic noise is simultaneously fit along with GW signals. These searches are computationally expensive owing to the large number of parameters in the model and can take weeks to run. In contrast, frequentist searches are significant faster to perform, although they use simplified signal models. Therefore, frequentist methods complement Bayesian methods and can be used to identify potential signals which can then be followed up with a full Bayesian search.

The $\mathcal{F}$ -statistic is a frequentist method to search for CWs that was initially developed for the Laser Interferometer Gravitational Wave Observatory (LIGO) [20]. The same approach was then applied to the Laser Interferometer Space Antenna (LISA) [21] and PTAs [22,23]. The $\mathcal{F}_p$ - and $\mathcal{F}_e$ -statistics, introduced in [24], are frequentist statistics for detecting continuous waves in PTA data. The $\mathcal{F}_p$ -statistic was used to analyze the NANOGrav 5-year and 11-year datasets [16,25], but has not been used to analyze subsequent datasets owing to the emergence of a common red noise signal. In the derivation of the $\mathcal{F}_p$ - and $\mathcal{F}_e$ -statistics, the authors assumed that only white noise and a CW signal were present, and did not consider the presence

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](#). Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

of either intrinsic red noise or a GWB. Recently, work has been done to use the $\mathcal{F}_e$ -statistic to analyze PTA data containing common red noise. The European Pulsar Timing Array (EPTA) used the $\mathcal{F}_e$ -statistic to search for CWs in their second data release in addition to a common process [26], and Furusawa *et al.* [27] looked at how the presence of a GWB affects the $\mathcal{F}_e$ -statistic.

In this paper, we look at the performance of the $\mathcal{F}_p$ -statistic in the presence of pulsar red noise and common red noise. In previous analyses [17,25,28], we used the maximum likelihood posterior parameters to calculate the $\mathcal{F}_p$ -statistic. However, in the case of the optimal statistic, a frequentist detection statistic for the GWB [29], it was found that noise marginalization leads to a more accurate distribution of recovered values [30]. In this paper, we investigate whether marginalizing over the red noise posteriors affects the $\mathcal{F}_p$ -statistic. We find that there is not a significant difference between the mean noise-marginalized $\mathcal{F}_p$ -statistic (NMFP) values and the maximum-likelihood $\mathcal{F}_p$ -statistic (MLFP) values. We use simulations to show that, in the absence of a CW signal, the NMFP values follow the expected chi-squared distribution for frequencies up to where the red noise is modeled, and in the presence of a CW the $\mathcal{F}_p$ -statistic can recover the signal at the correct frequency.

This paper is organized as follows. In Sec. II, we discuss our methods and software, including the implementation of the $\mathcal{F}_p$ -statistic and its differences from the $\mathcal{F}_e$ -statistic as described in [24]. We tested the $\mathcal{F}_p$ -statistic on simulated PTA datasets to study its performance in the presence of intrinsic pulsar red noise, common red noise, and a continuous wave; we present the results of our analyses on simulated datasets in Sec. III. In Sec. IV, we present the noise-marginalized results for the NANOGrav 12.5-year dataset [31], and compare these results to the results of Bayesian searches for individual binaries published in [17]. We give a brief summary of our work and our conclusions in Sec. V.

II. METHODS AND SOFTWARE

In this section, we review the signal model for a CW in PTA data. We assume the source is a circular binary evolving only due to GW emission. We also review the $\mathcal{F}_p$ -statistic, a frequentist statistic for detecting CWs. More details can be found in [24].

A. Signal model

GWs induce correlated changes in the pulse times arrival times for millisecond pulsars. The differences between the expected and observed pulse times of arrival are the *residuals*. We can write the residuals for each pulsar δt as [32]

$$\delta t = M\epsilon + Fa + n + s, \quad (1)$$

where M is the design matrix, which describes the linearized timing model, ϵ is a vector of the timing model parameter offsets, Fa describes red noise, n is a vector describing white noise, and s is a vector of the residuals induced by a CW.

The response of a pulsar with sky position (θ, ϕ) to a GW source is given by the antenna pattern functions [24],

$$F^+(\hat{\Omega}) = \frac{1}{2} \frac{(\hat{m} \cdot \hat{p})^2 - (\hat{n} \cdot \hat{p})^2}{1 + \hat{\Omega} \cdot \hat{p}}, \quad (2)$$

$$F^\times(\hat{\Omega}) = \frac{(\hat{m} \cdot \hat{p})(\hat{n} \cdot \hat{p})}{1 + \hat{\Omega} \cdot \hat{p}}, \quad (3)$$

where

$$\hat{m} = -\sin \phi \hat{x} + \cos \phi \hat{y}, \quad (4)$$

$$\hat{n} = -\cos \theta \cos \phi \hat{x} - \cos \theta \sin \phi \hat{y} + \sin \theta \hat{z}, \quad (5)$$

and $\hat{\Omega}$ is a unit vector pointing from the GW source to the Solar System barycenter. The induced residuals can be written,

$$s(t, \hat{\Omega}) = F^+(\hat{\Omega})\Delta s_+(t) + F^\times(\hat{\Omega})\Delta s_\times(t), \quad (6)$$

where $\Delta s_{+, \times}$ is the difference between the signal induced at the Earth and at the pulsar (the so-called “Earth term” and “pulsar term”),

$$\Delta s_{+, \times}(t) = s_{+, \times}(t_p) - s_{+, \times}(t). \quad (7)$$

The time at which the GW passes the Solar System barycenter t and the time at which it passes the pulsar t_p are related by

$$t_p = t - L(1 + \hat{\Omega} \cdot \hat{p}), \quad (8)$$

where L is the distance to the pulsar and $\hat{p}$ is a unit vector pointing from the Earth to the pulsar.

For a circular binary the signal is given by [33–35]

$$s_+(t) = \frac{\mathcal{M}^{5/3}}{d_L \omega(t)^{1/3}} [-\sin 2(\Phi(t) - \Phi_0)(1 + \cos^2 i) \cos 2\psi - 2 \cos 2(\Phi(t) - \Phi_0) \cos i \sin 2\psi], \quad (9)$$

$$s_\times(t) = \frac{\mathcal{M}^{5/3}}{d_L \omega(t)^{1/3}} [-\sin 2(\Phi(t) - \Phi_0)(1 + \cos^2 i) \sin 2\psi + 2 \cos 2(\Phi(t) - \Phi_0) \cos i \cos 2\psi], \quad (10)$$

where ω is the orbital frequency, i is the inclination angle of the SMBHB, ψ is the GW polarization angle, d_L is the

luminosity distance to the source, and $\mathcal{M} \equiv (m_1 m_2)^{3/5} / (m_1 + m_2)^{1/5}$ is a combination of the black hole masses m_1 and m_2 called the “chirp mass”. Note that the variables $\mathcal{M}$ and ω are the observed redshifted values, which are related to the rest-frame values $\mathcal{M}_r$ and ω_r according to $\mathcal{M}_r = \mathcal{M}/(1+z)$ and $\omega_r = \omega(1+z)$. Currently PTAs are only sensitive to sources in the local Universe for which $(1+z) \approx 1$.

As the binary emits GWs, the orbital frequency of the binary evolves according to

$$\frac{d\omega}{dt} = \frac{96}{5} \mathcal{M}^{5/3} \omega^{11/3}. \quad (11)$$

For typical SMBHBs emitting in the PTA band, $d\omega/dt$ is not significant over the timescale of the PTA dataset (~ 20 yrs), but can be over the light-travel time between the pulsars and the Earth (~ 1000 yrs). Then the orbital frequency at the pulsar $\omega_p \equiv \omega(t_p)$ is related to the orbital frequency at the Earth $\omega_0 \equiv \omega(t)$ according to

$$\omega_p = \omega_0 \left[1 + \frac{256}{5} \mathcal{M}^{5/3} \omega_0^{8/3} L (1 + \hat{\Omega} \cdot \hat{p}) \right]^{-3/8}. \quad (12)$$

For slowly evolving sources, we can approximate this as

$$\omega_p \approx \omega_0 \left[1 - \frac{96}{5} \mathcal{M}^{5/3} \omega_0^{8/3} L (1 + \hat{\Omega} \cdot \hat{p}) \right]. \quad (13)$$

B. $\mathcal{F}_p$ -statistic

The likelihood of a continuous wave being present is given by

$$p(\delta t|s) = \frac{\exp[-\frac{1}{2}(\delta t - s)^T C^{-1}(\delta t - s)]}{\sqrt{\det[2\pi C]}}, \quad (14)$$

where C is the total covariance matrix of the PTA. For signals uncorrelated between pulsars the covariance matrix becomes a block diagonal matrix of the individual pulsar autocovariances, i.e., $C = \text{diag}(\Sigma_1, \Sigma_2, \dots, \Sigma_N)$, where Σ_α is

$$\Sigma_\alpha = N_\alpha + T_\alpha^T B_\alpha T_\alpha, \quad (15)$$

where T is a block matrix containing the design matrix M and the Fourier matrix F , N is the white noise covariance matrix, and B is a block diagonal matrix containing the covariances of the priors of parameters ϵ and a . An important note is that in the presence of a common correlated signal, the total covariance matrix C is now a dense block matrix with pulsar autocovariances Σ_α in the diagonal terms and cross-covariances Σ_{ab} , defined by

$$\Sigma_{\alpha\beta} = \Gamma_{\alpha\beta} F \phi_{\text{GWB}} F^T, \quad (16)$$

where F is the Fourier matrix, ϕ_{GWB} is the power spectral density of the correlated background, and Γ_{ab} is the Hellings-Downs coefficient between pulsars a and b [36]. In the case of a CURN, the $T^T B T$ matrix contains intrinsic red noise, and a common red noise process for all pulsars in Eq. (15). In the case of a correlated common red noise process (the GWB), the total covariance matrix C now contains off diagonal terms that follow Eq. (16).

We start with the log-likelihood ratio of a continuous signal s being present within our residuals δt to no signal,

$$\ln \Lambda = \ln \frac{p(\delta t|s)}{p(\delta t|0)}, \quad (17)$$

$$\ln \Lambda = \sum_\alpha (\delta t_\alpha | \tilde{s}_\alpha) - \frac{1}{2} (\tilde{s}_\alpha | \tilde{s}_\alpha). \quad (18)$$

Here, the terms in the parentheses are the inner product with the inverse covariance matrix, i.e., $(x|y) = x \Sigma^{-1} y$, where Σ is the pulsar autocovariance matrix.

The signal s from a continuous wave source is represented as the sum of sines and cosines for the α pulsars,¹

$$s_\alpha(t, \hat{\Omega}) = \sum_{i=1}^2 b_{i\alpha}(\zeta, \iota, \psi, \Phi_0, \phi_\alpha, \theta, \phi) B_\alpha^i(t, \omega_0), \quad (19)$$

where $\zeta = \mathcal{M}^{5/3}/d_L$ is the amplitude, ι is the inclination, ψ is the polarization, Φ_0 is the gravitational wave initial phase, $\phi_\alpha = \omega L_\alpha (1 + \hat{\Omega} \cdot \hat{p}) + \Phi_0$ is the pulsar dependent phase, and θ, ϕ are the sky position for the source. The only intrinsic parameter is ω_0 . There are two terms (sine and cosine) for the amplitudes $b_{i\alpha}$ and the basis functions B_α^i .

The amplitudes $b_{1\alpha}$ and $b_{2\alpha}$ are given by

$$b_{1\alpha} = \zeta [(1 + \cos^2 \iota) (F_\alpha^+ \cos 2\psi + F_\alpha^\times \sin 2\psi) \times (\cos \Phi_0 - \cos \phi_\alpha) + 2 \cos \iota (F_\alpha^+ \sin 2\psi - F_\alpha^\times \cos 2\psi) (\sin \Phi_0 - \sin \phi_\alpha)], \quad (20)$$

$$b_{2\alpha} = -\zeta [(1 + \cos^2 \iota) (F_\alpha^+ \cos 2\psi + F_\alpha^\times \sin 2\psi) \times (\sin \Phi_0 - \sin \phi_\alpha) - 2 \cos \iota (F_\alpha^+ \sin 2\psi - F_\alpha^\times \cos 2\psi) (\cos \Phi_0 - \cos \phi_\alpha)], \quad (21)$$

and the basis functions are simply,

$$B_\alpha^1(t) = \frac{1}{\omega_0^{1/3}} \sin 2\omega_0 t, \quad (22)$$

¹Here we are only considering circular SMBBHs. For a discussion of frequentist searches for eccentric SMBBHs, see Ref. [37].

$$B_a^2(t) = \frac{1}{\omega_0^{1/3}} \cos 2\omega_0 t. \quad (23)$$

The log-likelihood then becomes

$$\ln \Lambda = \sum_{\alpha=1}^M \left[b_{i\alpha}(r_\alpha | B_\alpha^i) - \frac{1}{2} (B_\alpha^i | B_\alpha^j) b_{i\alpha} b_{j\alpha} \right]. \quad (24)$$

Here, the indices i, j are for the sine and cosine terms and M is the total number of pulsars. We can represent the inner products as matrices $P_\alpha^i = (r_\alpha | B_\alpha^i)$ and $Q_\alpha^{ij} = (B_\alpha^i | B_\alpha^j)$, and maximize the log-likelihood to get the relationship $b_{i\alpha} = Q_{ij}^\alpha P_\alpha^j$, where $Q_{ij}^\alpha = (Q_\alpha^{ij})^{-1}$. The matrices P_α^i has dimensions of $N_{\text{TOA},\alpha} \times 2$, and Q_α^{ij} is a 2×2 matrix.

We maximize the extrinsic parameters $b_{k\alpha}(\zeta, \iota, \psi, \Phi_0, \phi_\alpha, \theta, \phi)$, and only have one intrinsic parameter which is our frequency ω as in [24]. Thus we get the maximized log-likelihood and the $\mathcal{F}_p$ statistic,

$$\ln \Lambda = \frac{1}{2} \sum_{\alpha=1}^M P_\alpha^i Q_{ij}^\alpha P_\alpha^j, \quad (25)$$

$$2\mathcal{F}_p = \sum_{\alpha=1}^M P_\alpha^i Q_{ij}^\alpha P_\alpha^j. \quad (26)$$

Here, the distribution of the $2\mathcal{F}_p$ -statistic is a chi-squared with degrees of freedom $2N$, where N is the number of pulsars, and a noncentrality of $\rho^2 = (\tilde{s} | \tilde{s})$, in the case of a continuous wave, where ρ is the signal-to-noise (S/N) ratio of the continuous wave. Each frequency bin will follow the standard chi-squared distribution, except for bins with a continuous wave, which will follow a noncentral chi-squared distribution. Since there are six independent extrinsic global variables $(\zeta, \iota, \psi, \Phi_0, \theta, \phi)$, and one extrinsic variable per pulsar (ϕ_α) , we have a total of $6 + N$ unknown variables and $2N$ equations, where N is the number of pulsars. Therefore, we need a minimum of six pulsars to solve for all variables [24].

For our simulations based on the NANOGrav 15-year dataset, there are 67 pulsars, so the expected distribution of the $\mathcal{F}_p$ -statistic if there is only white noise present is a chi-squared distribution with 134 degrees of freedom, as shown in Fig. 1. To determine the presence of a continuous wave, we compute the False Alarm Probability (FAP) by comparing the measured value of the $\mathcal{F}_p$ -statistic to the null distribution [20,24,38]. The probability distributions for when the signal is absent (p_0) and present (p_1), and intrinsic parameters are known are

$$p_0(\mathcal{F}) = \frac{\mathcal{F}^{n/2-1}}{(n/2-1)!} \exp(-\mathcal{F}), \quad (27)$$

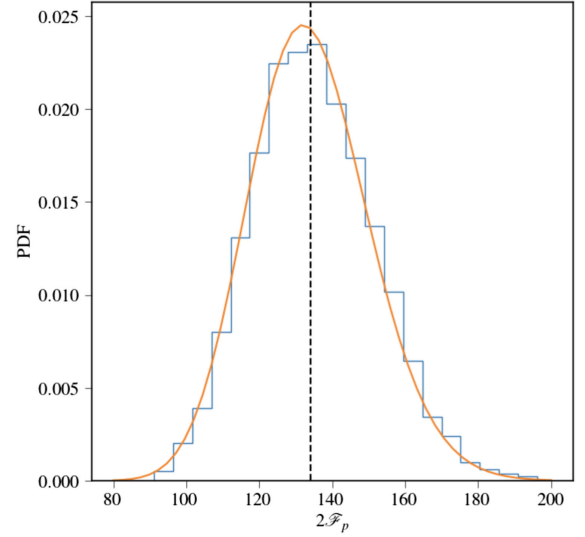


FIG. 1. The blue histogram shows the $2\mathcal{F}_p$ values for 10 simulations of the 15 yr dataset with only white noise injected, marginalized across 150 frequency bins. The orange curve is the expected chi-squared distribution with zero noncentrality and the black dashed line is the expected degrees of freedom.

$$p_1(\mathcal{F}, \rho) = \frac{(2\mathcal{F})^{(n/4-1/2)}}{\rho^{(n/2-1)}} I_{n/2-1}(\rho\sqrt{2\mathcal{F}}) \times \exp\left(-\mathcal{F} - \frac{1}{2}\rho^2\right), \quad (28)$$

where n is the degrees of freedom ($n = 2M$), ρ is the signal-to-noise ratio, and I_x is the modified Bessel function of the first kind. The probability that $\mathcal{F}$ exceeds a threshold $\mathcal{F}_0$ when no signal is present is the FAP,

$$P_F(\mathcal{F}_0) = \int_{\mathcal{F}_0}^{\infty} p_0(\mathcal{F}) d\mathcal{F} = \exp(-\mathcal{F}_0) \sum_{k=0}^{n/2-1} \frac{\mathcal{F}_0^k}{k!}. \quad (29)$$

Since we do not know the intrinsic parameters we must adjust the FAP values to reflect the number of independent cells N_c , as shown in [24,38],

$$P_F^T(\mathcal{F}_0) = 1 - [1 - P_F(\mathcal{F}_0)]^{N_c}, \quad (30)$$

where in our analysis, N_c is the number of frequency bins over which we calculate the $\mathcal{F}_p$ value. This value gives us the probability the $\mathcal{F}$ exceeds $\mathcal{F}_0$ in one or more bins.

Since we maximize over the extrinsic parameters, which include the sky position (θ, ϕ) , the $\mathcal{F}_p$ -statistic is an incoherent search which marginalizes over the sky location of the source and sums the power across all pulsars. In this respect, it is analogous to Bayesian all-sky searches. In contrast, the $\mathcal{F}_e$ -statistic is a coherent search, and the sky location of the source and GW frequency are free parameters. In this paper, we only discuss the $\mathcal{F}_p$ -statistic; other

papers have examined the $\mathcal{F}_e$ -statistic in the presence of pulsar red noise and common red noise [26,27].

C. Software

All of the results below utilize our new package FASTFP² which is publicly available on Github. Both the original $\mathcal{F}_p$ -statistic as well as the NMFP described above are implemented in our code. It is programmed entirely using the JAX [39] PYTHON library which leads to roughly an order of magnitude speedup for a single $\mathcal{F}_p$ evaluation relative to the code utilized in frequentist CW analyses of previous NANOGrav datasets. This framework also supports efficient batched evaluation, allowing for simultaneous calculation of thousands of $\mathcal{F}_p$ values and making noise marginalization techniques computationally feasible. Additionally our code natively supports GPU acceleration, further reducing overall runtime.

We use the standard suite of PTA analysis software to carry our simulated data studies and Bayesian analyses. We use the libstempo package to create the three different configurations of simulated datasets. We use enterprise [40] and enterprise_extensions [41] to set up the PTA model and calculate the likelihood, and we use PTMCMCSampler [42] to perform the Monte Carlo Markov Chain sampling.

III. RESULTS: SIMULATED DATASETS

We use simulated datasets to test the performance of the $\mathcal{F}_p$ -statistic in the presence of pulsar noise, common uncorrelated red noise (CURN), a GWB, and a CW signal. We base our datasets on the NANOGrav 15-year dataset [43]; we simulate 67 pulsars observed for up to 15.9 years, with sky positions and timing model parameters taken from the NANOGrav 15-year dataset. We include white noise and intrinsic red noise in each pulsar; the injected red noise parameters are given in Table I. For the simulations including CURN or a GWB, we model the power spectrum as a power law with amplitude and spectral index given in Table I.

For the CW injection we use the most sensitive sky position as given by [18] with parameters given in Table II. We considered two CW injections, one at 8 nHz and one at 16 nHz, with S/N ratio of 5.4 and 9.84, respectively, for cases with a medium and high S/N ratio. The signal to noise ratio can be calculated using Eq. (18),

$$S/N = \sqrt{\sum_{\alpha} (\tilde{s}_{\alpha} | \tilde{s}_{\alpha})}, \quad (31)$$

in the presence of a CW, the noncentrality parameter should be ρ^2 , where ρ is the S/N ratio of the signal.

²<https://github.com/gabefreedman/fastfp>.

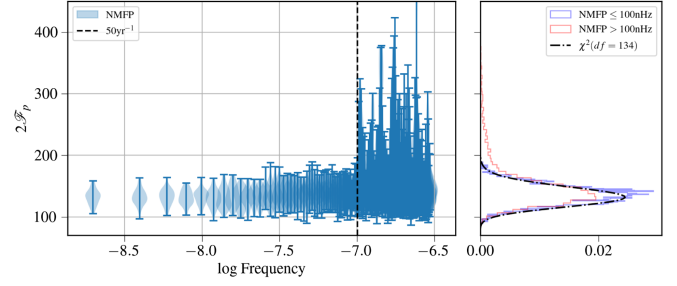


FIG. 2. (Left) The $2\mathcal{F}_p$ value for simulations with white noise and intrinsic red noise for the NMFP values. Both the Bayesian and the frequentist analysis were performed for 50 red noise frequency components (black dashed line), beyond which the $\mathcal{F}_p$ value becomes large and unreliable. (Right) The histograms showing the $\mathcal{F}_p$ values below 100 nHz (50 yr^{-1}) $\mathcal{F}_p$ in light blue, and the $\mathcal{F}_p$ values above 100 nHz in dark blue.

For datasets containing red noise, we first run a Bayesian analysis to calculate the posteriors of all red noise parameters. Our Bayesian analysis only takes into consideration the autocorrelation terms and takes hours to days to run, compared to a full Bayesian analysis with cross-correlations which can take days to weeks. We then compute the $\mathcal{F}_p$ -statistic in two ways: (1) Compute the $\mathcal{F}_p$ -statistic with the red noise parameters set to their maximum-likelihood values, as was done in [16,24]; (2) Marginalizing over the red noise using the red noise posteriors. In both methods, an initial Bayesian analyses is necessary to get posteriors for the red noise parameters. The posteriors of this Bayesian analysis can then be used for $\mathcal{F}_p$ -statistic calculations at all frequencies. The initial Bayesian analysis also takes significantly less time than a full Bayesian analysis including red noise and a CW.

This type of noise marginalization has been done in frequentist searches for the GWB. As shown in [30,44], marginalizing over the red noise more accurately recovers the amplitude and significance of the GWB than using maximum-likelihood noise parameters. However, in our analysis of the $\mathcal{F}_p$ -statistic, we found little difference between these two methods, as we show below.

To quantify the significance of a detection, compute the FAP value following Eq. (30). We adopt a threshold of $P_F(\mathcal{F}_0) \leq 10^{-4}$ for a detection, as was done in the NANOGrav 11-yr continuous wave search [16].

A. White noise and intrinsic red noise

We considered datasets with both white noise and intrinsic red noise. Figure 2 shows the recovered values of the $2\mathcal{F}_p$ -statistic between 2–300 nHz. The recovered values follow the expected chi-squared distribution below 100 nHz; however, above that the recovered values no longer follow the expected distribution. This transition occurs at the frequency where we truncate the model for the pulsar's intrinsic red noise, demonstrating the importance

of modeling pulsar noise when searching for CWs. For the remainder of the paper, we limit the $\mathcal{F}_p$ -statistic analyses to below 100 nHz. In Appendix C we show the percentile-percentile plots for both the mean NMFP and MLFP across all simulations for each of the modeled frequencies. We calculate the percentage with respect to the expected chi-squared distributions.

B. White noise, intrinsic red noise, and CURN

Next we applied the $\mathcal{F}_p$ -statistic to datasets containing CURN in addition to pulsar white noise and intrinsic red noise. When we neglected to model the CURN, we found very high values of $2\mathcal{F}_p$ at the lowest frequencies, indicating confusion between CURN and a CW, but at higher frequencies $2\mathcal{F}_p$ followed a chi-squared distribution with noncentrality parameter of zero. This is because the CURN signal has a steep power spectrum, and so it only affects the $\mathcal{F}_p$ -statistic analysis at low frequencies.

When we included the CURN in our model, as described in Sec. II B, we again found that $2\mathcal{F}_p$ followed a chi-squared distribution with noncentrality parameter of zero, indicating no detection of a CW. This is consistent with the findings of [27], where it was found that if the background is modeled as noise, it can be distinguished from the continuous wave. In the top panel (blue) of Fig. 3 we see that the means of the NMFP values are in agreement the expected chi-squared distribution across all 50 frequencies when we marginalize over the intrinsic red noise as well as the CURN parameters. In the middle panel (red) we show the recovery when the CURN parameters are fixed to the maximum likelihood posteriors, and only marginalize the intrinsic red noise parameters.

We find good agreement between the $2\mathcal{F}_p$ values when we marginalize over the CURN parameters or use the maximum likelihood CURN parameters, although there is larger variance in the $2\mathcal{F}_p$ values when marginalizing over CURN at very low frequencies.

C. White noise + Intrinsic Red Noise + HD simulations

In our previous set of simulations, we considered an uncorrelated common process, the common process was uncorrelated, but a GWB also includes the presence of interpulsar correlations described by the Hellings-Downs (HD) coefficients [36]. To test the effect of correlations, we created a set of simulations with white noise, intrinsic red noise, and a realistic background with HD correlations. When running our Bayesian analysis, however, we modeled the common process as CURN rather than a correlated process due to the computationally intense process of inverting a dense covariance matrix described in Sec. II B. This allowed us to examine whether not including the cross-correlations in our model would lead to biased results in the CW search.

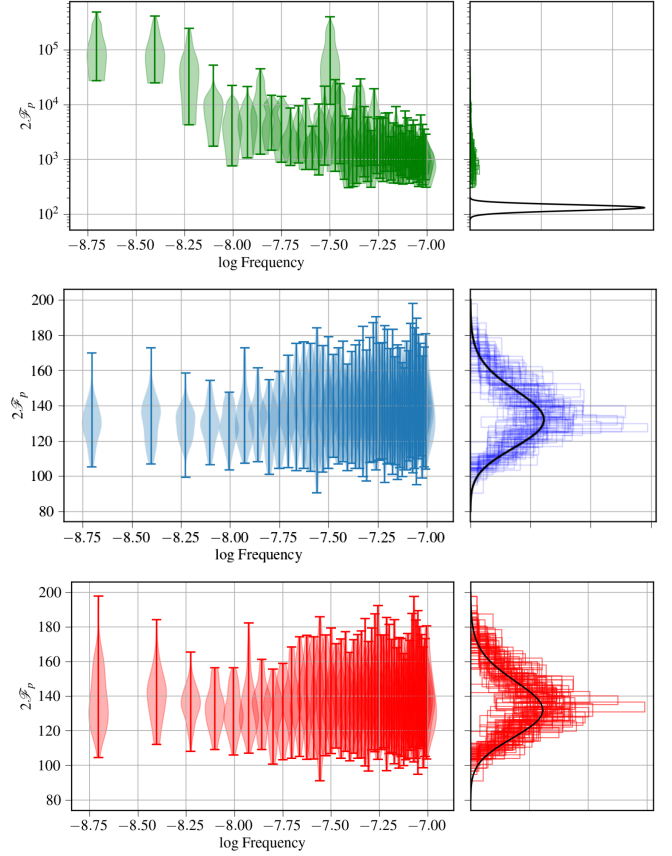


FIG. 3. TOP: (Left) The violin plot and histograms for $\mathcal{F}_p$ values when we do not include the CURN model. We see that at all frequencies the $2\mathcal{F}_p$ values are greatly overestimated. (Right) The histograms for the $2\mathcal{F}_p$ values as compared to the expected central chi-squared distribution with 134 degrees of freedom. MIDDLE: The violin plots for the mean NMFP values of 50 simulations with white noise, intrinsic red noise, and a common process as a function of frequency. BOTTOM: The figure shows the violin plots for the NMFP values when the common process is kept fixed to the maximum likelihood values and only intrinsic red noise parameters are marginalized over. At low values, the $2\mathcal{F}_p$ values recover a slightly higher mean and variance due to the common process not being modeled correctly. The frequency ranges of all three plots are ~ 2 –100 nHz.

We plot the $2\mathcal{F}_p$ values in Fig. 4 and find that the recovered values do not vary from the CURN results. This is due to the fact that the implementation of the $\mathcal{F}_p$ -statistic only takes into account pulsar autocorrelations and does not consider the cross-correlations as shown in Eq. (25). In the bottom panel of Fig. 4 we also show the S/N ratio of the injected correlations, which range from 2 to more than 12. We conclude that modeling the GWB as CURN does not significantly affect the null distribution of the $\mathcal{F}_p$ -statistic when the cross-correlations of the GWB are at this level of significance.

It is important to note that both the GWB and the CW induce cross-correlations between pairs of pulsars, however

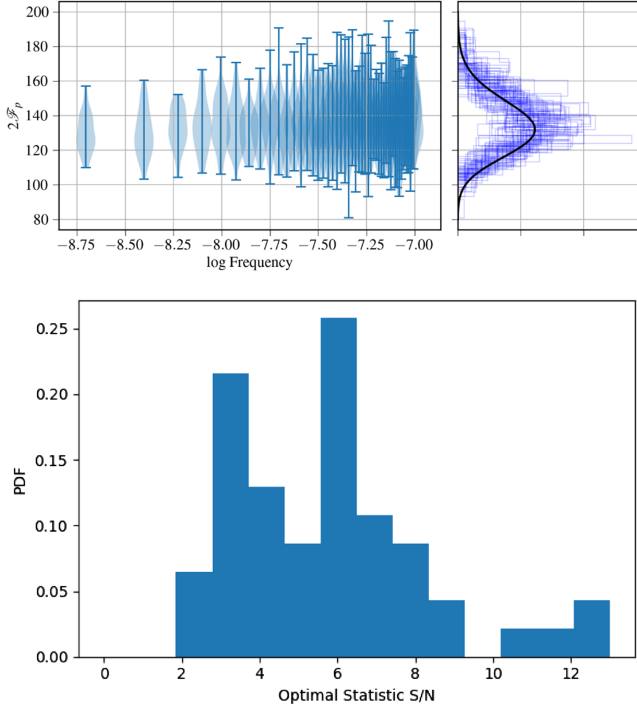


FIG. 4. TOP: The violin plots of the NMFP values for simulations with white noise, intrinsic red noise, and a common correlated background (left) and a plot of the histograms of $\mathcal{F}_p$ values for each frequency (right). We see that this figure is nearly identical to Fig. 3. BOTTOM: Distribution of the optimal statistic S/N ratio for the simulated datasets. For the simulations used here, the S/N ratio of the cross-correlations ranged from 2 to more than 12.

using the dense covariance matrix for the $\mathcal{F}_p$ -statistic, as described in Sec. II B, is computationally expensive and will hinder any advantage of the speed we achieve with GPU acceleration. When considering a strong background, the effect of the cross correlations may become significant, however at our current level of significance this is not an issue. The inclusion of the full auto and cross-covariance matrices in calculating the $\mathcal{F}_p$ -statistic is beyond the scope of this paper, but should be revisited in future studies.

D. Continuous wave in the presence of CURN

Finally, we look at how well the $\mathcal{F}_p$ -statistic can recover CW signals in the presence of white noise, intrinsic red noise, and CURN. We used the parameters in Table II, which place the source at the most sensitive sky location for the NANOGrav 15-year dataset [18]. We kept the chirp mass fixed and considered two different GW frequencies, 8 nHz and 16 nHz as well as two slightly different luminosity distances, which corresponded to S/N ratios of 5.4 and 9.1, respectively.

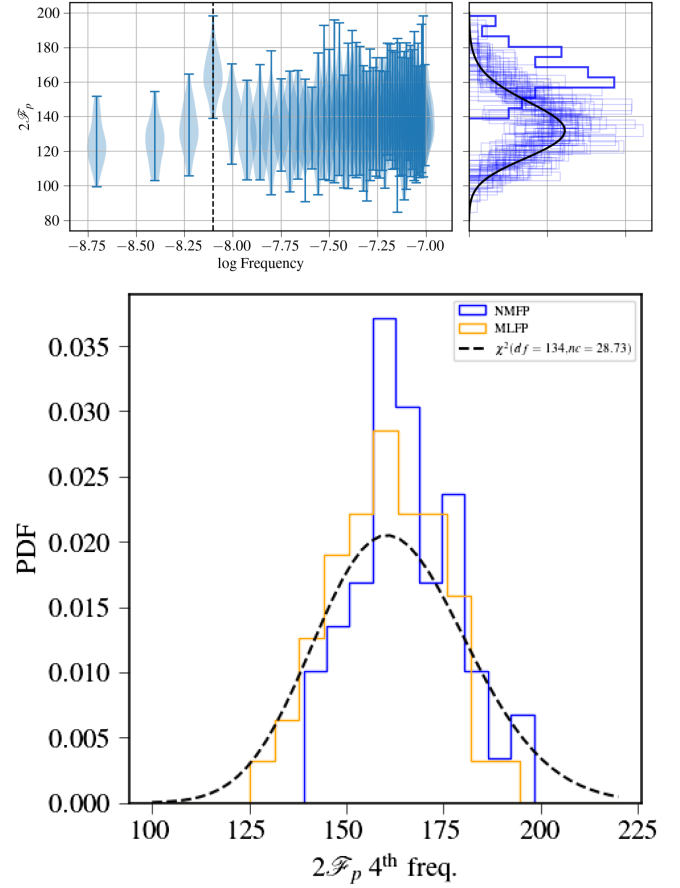


FIG. 5. (Left) The violin plots for both the NMFP values of simulations with white noise, intrinsic red noise, a common process, and a continuous wave signal with a medium S/N ratio across frequencies ~ 2 –100 nHz. (Right) Histograms of all frequencies compared to the null distribution (black curve). The darker colored histograms indicate the fourth frequency component. Bottom figure: Histogram for the NMFP (blue) and MLFP (orange) values at the injected frequency of 8 nHz compared to the expected noncentral chi-squared distribution.

1. 8 nHz injection (S/N=5.4)

Figure 5 shows the violin plots of $2\mathcal{F}_p$ for a CW injection at 8 nHz. We see that at the fourth frequency bin, which corresponds to the frequency of the injected CW, $2\mathcal{F}_p$ does not follow a chi-squared distribution with $\rho^2 = 0$, but it does at the other frequencies. In the bottom panel of Fig. 5 we show the histogram of the NMFP and MLFP values at the injected frequency of 8 nHz. We show that the recovered values follow the expected chi-squared distribution with a noncentrality of 28.73. Figure 6 shows a histogram of the p -values for $2\mathcal{F}_p$ at 8 nHz. We find that at 8 nHz 56% of simulations have $p < 0.05$. In contrast, at every other frequency, we find $p < 0.05$ in 9.33% of simulations. From this we conclude that, although the $\mathcal{F}_p$ -statistic values at 8 nHz follow a different distribution

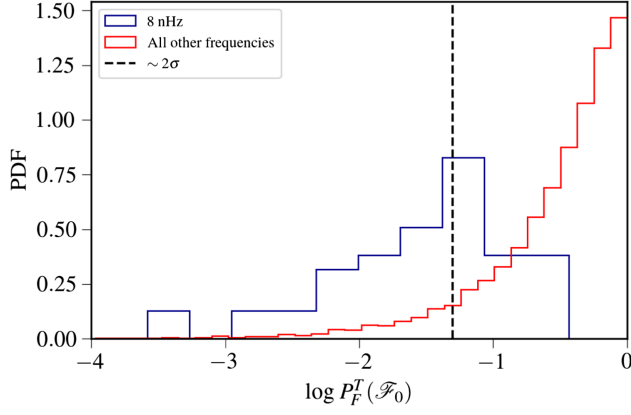


FIG. 6. The log FAP values for the injected frequency of 8 nHz (dark blue) and all other frequencies (red). The dashed-line indicates a p-value of 0.05, corresponding to $\sim 2\sigma$. At the injected frequency, 56% of NMFP values have an FAP value less than 0.05. None of the simulations fall below the 11 yr threshold of 10^{-4} .

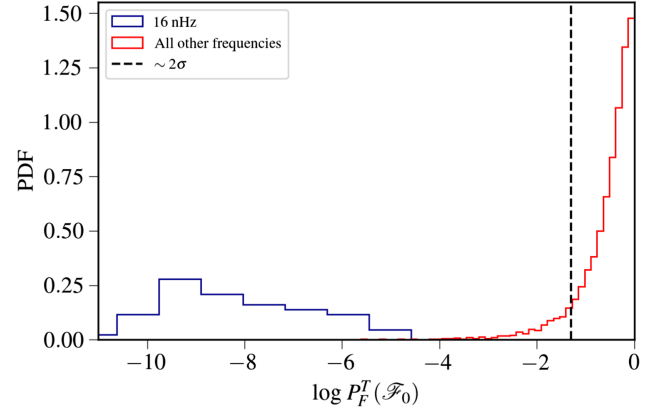


FIG. 8. Histogram of log FAP values for the injected frequency at 16 nHz (dark blue) and all other frequencies (light blue). The dashed line indicates a p-value of 0.05, corresponding to $\sim 2\sigma$. All simulations have $p < 10^{-4}$ at the injected frequency.

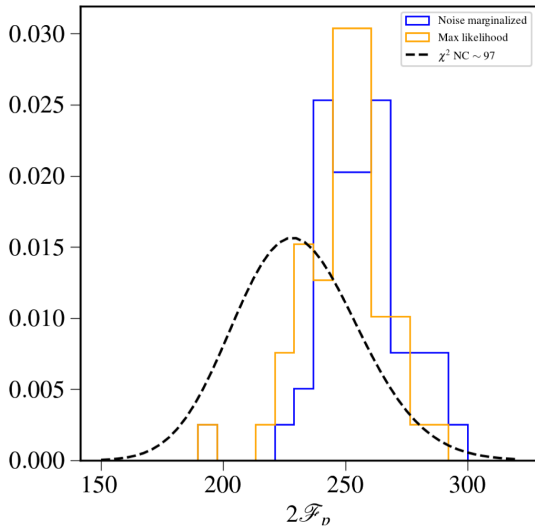
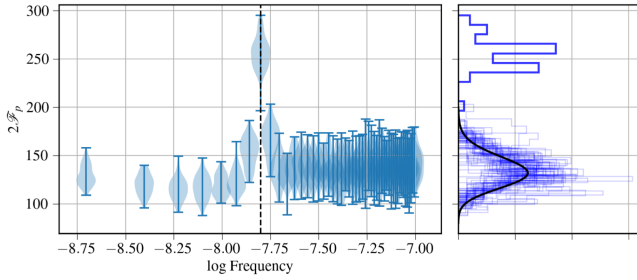


FIG. 7. Top figure: (Left) The violin plots for the mean NMFP values of 50 simulations with white noise, intrinsic red noise, a common process, and a high S/N ratio continuous wave as a function of frequency across frequencies ~ 2 –100 nHz. (Right) The histograms for the $2\mathcal{F}_p$ values as compared to the expected central chi-squared distribution with 134 degrees of freedom. Bottom figure: The histograms for the $\mathcal{F}_p$ -statistic at 16 nHz compared to the expected chi-squared distribution with noncentrality of 97.

than at the other frequencies, the significance is too low to be considered a detection.

2. 16 nHz injection ($S/N = 9.84$)

Figure 7 shows the violin plots of the $2\mathcal{F}_p$ values. We see that at the eighth frequency, which corresponds to the frequency of the CW injection, $2\mathcal{F}_p$ does not follow a chi-squared distribution with noncentrality parameter of zero. There is also a shift in the distribution of $2\mathcal{F}_p$ at the frequencies near the CW injection, especially the ninth frequency (~ 18 nHz). At all other frequencies however, the distributions follows the expected central chi-squared distribution. In the bottom panel of Fig. 7 we see that the recovered NMFP and MLFP values at the injected frequency follow a distribution with a noncentrality that is slightly higher than the expected noncentrality of 97. This is likely because the S/N ratio is extremely sensitive to small changes in our parameters and the associated uncertainty with our simulations could cause a slight bias in S/N ratio. However, our results are consistent with the presence of a CW.

In Fig. 8, we show the FAP values calculated from the mean NMFP values for 50 simulations at 16 nHz. We find that all of the simulations have $p < 0.05$, and 88% of simulations have $p < 10^{-4}$ at the injected frequency of 16 nHz. In contrast, across the other frequencies we find $p < 0.05$ in 9.3% of simulations, and $p < 10^{-4}$ in 0.3% of simulations. From this we conclude that the $\mathcal{F}_p$ -statistic is able to detect a CW at the injected frequency of 16 nHz in the majority of simulations.

IV. RESULTS: NANOGrav 12.5-YEAR DATASET

In this section we present the $\mathcal{F}_p$ values of the NANOGrav 12.5 yr dataset of 45 pulsars with noise marginalization to complement the published Bayesian

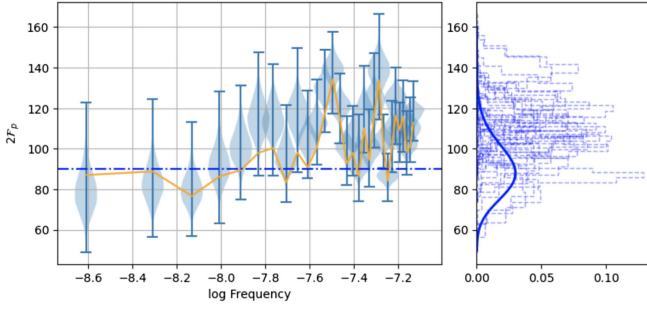


FIG. 9. The $2\mathcal{F}_p$ values for the 12.5 yr dataset for frequencies ranging from ~ 2.5 – 76 nHz. On the left panel, the blue violin plots are the NMFP values for 20,000 red noise parameters. The orange line is the noise maximized $\mathcal{F}_p$ values. The dot-dashed blue line is the expected mean for a chi-squared distribution with 90 degrees of freedom. In the right panel, the dark blue histogram is the null chi-squared distribution with 90 degrees of freedom and zero noncentrality.

CW search results [17]. We use the same noise models that were used in [17] and other analyses of the NANOGrav 12.5-year dataset. The intrinsic red noise was modeled with 30 frequency components, and the common process was modeled as CURN with five frequency components. We also include white noise correlated among different epochs (referred to as ECORR), which we did not include in our simulations. Since this white noise is correlated at different observing epochs it is possible that it results in a slightly different distribution in $\mathcal{F}_p$ values than the null distribution.

In Fig. 9 we show the violin plot of the noise-marginalized $\mathcal{F}_p$ values for the 12.5 yr dataset. We see that the first few frequencies are within the expected range of the null hypothesis. After the first five frequency components, where we stop modeling the common process, we get higher $\mathcal{F}_p$ values. At $f_{\text{gw}} = 1 \text{ yr}^{-1} = 31.9$ nHz, we recover higher values due to the insensitivity of PTAs to GW's at this frequency owing to the Earth's orbital period, but these values are still consistent with the null distribution.

In Fig. 10 we plot the mean FAP values of the calculated NMFP values for the 12.5 yr dataset. In the bottom panel of Fig. 10, we show the distribution of FAP values at 51.5 nHz, which has the highest $\mathcal{F}_p$ -value. At this frequency, the mean FAP value is 0.04. Taking into account the trials factor from searching over 30 frequencies, we conclude that these results are consistent with only noise being present in the data; the p -value with the Bonferroni correction is $p = 0.04 \times 30 \sim 1$.

In this analysis, we have not incorporated any advance noise modeling in order to use the same noise models as were used in the 12.5-year Bayesian CW search. We note that Bayesian searches for CWs have found that unmodeled

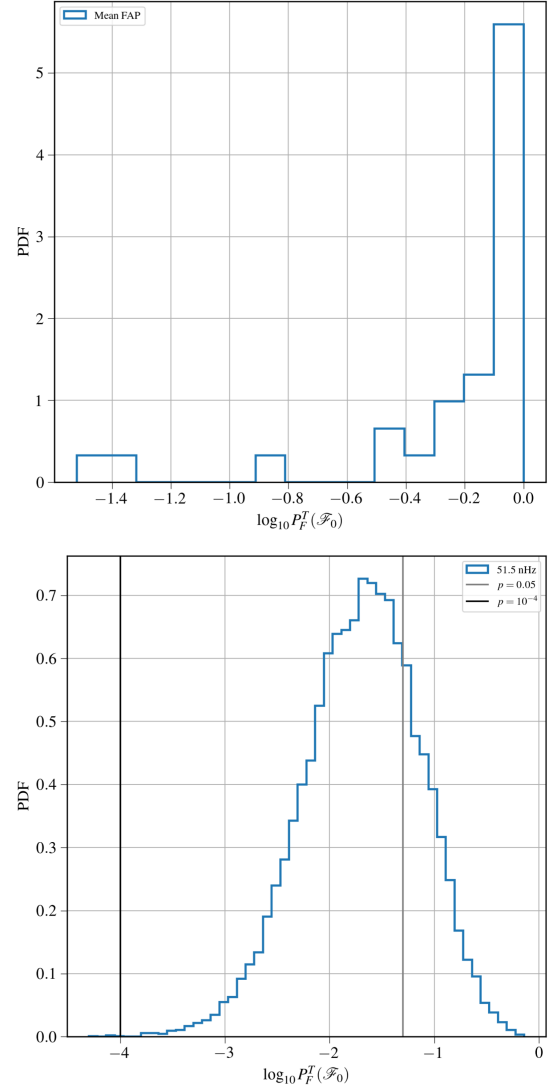


FIG. 10. (TOP) Mean FAP values, marginalized over noise posteriors, for each frequency in the 12.5 yr dataset. None of the frequencies have a mean FAP below 10^{-4} , and the lowest mean FAP value occurs at 51.5 nHz, with a value of 0.04. (BOTTOM) The full distribution of FAP values at 51.5 nHz over noise parameters. We see that more than half of the values fall below the 0.05 threshold, and only 0.02% fall below the threshold of 10^{-4} . Therefore, we conclude that there is no evidence of a CW in the 12.5 yr dataset. This result is consistent with the Bayesian results published in [17].

noise in individual pulsars has a significant effect on Bayes factors [16–18]. Significant work is ongoing developing more advanced noise models that incorporate chromatic noise and the Solar wind with the goal of improving GW searches by reducing confusion between pulsar noise and GW signals [45–48]. These models can be used in combination with the $\mathcal{F}_p$ -statistic to do frequentist CW searches.

V. CONCLUSIONS

In this paper, we used simulations to study the performance of the $\mathcal{F}_p$ -statistic in the presence of pulsar red noise and common red noise consistent with a GWB. We introduced noise-marginalization to the $\mathcal{F}_p$ -statistic and compared it to the previously used $\mathcal{F}_p$ values calculated from the maximum likelihood noise parameters. We found that the NMFP and MLFP give similar results to each other and are able to recover accurate values in the presence of white noise and intrinsic pulsar red noise.

We were also able to recover the expected $\mathcal{F}_p$ values even in the presence of white noise, intrinsic red noise, and a common process, with or without correlations. This is due to the fact that the $\mathcal{F}_p$ -statistic is only concerned with pulsar autocorrelations, and not cross-correlations. At our current level of significance, the autocorrelations are larger compared to our cross-correlations and so we are able to recover the same null distribution for both a CURN and a GWB. In previous analyses of the $\mathcal{F}_p$ -statistic, models only included white noise, or white noise and a common uncorrelated process. We have shown that the NMFP can accurately retrieve values for more datasets with more realistic noise descriptions. We also showed that using the maximum likelihood GWB parameters is able to recover comparable $\mathcal{F}_p$ values to marginalizing over GWB parameters.

We showed that both the NMFP and MLFP were comparable even in the presence of a continuous wave with medium or high S/N ratio. We found that for the medium S/N ratio case, the recovered $\mathcal{F}_p$ values followed a chi-squared distribution with the expected noncentrality. In the case of the high S/N ratio simulations, we found that both the MLFP and NMFP recovered $\mathcal{F}_p$ values that followed a χ^2 distribution with a larger noncentrality than expected, but were still consistent with a detection.

For both sets of continuous wave simulations, we measured the significance of the detection using the FAP values, with thresholds of 0.05 and 10^{-4} . We showed that in the case of the medium S/N ratio more than half of simulations recovered a value less than 0.5 at the injected frequency. For the high S/N ratio case, all the values at the injected frequency fell below the 0.05 threshold and nearly 90% of them fell below the 10^{-4} threshold.

We also reanalyzed the NANOGrav 12.5 yr dataset using the NMFP and compared it to the MLFP values. We found no evidence of a CW in our analysis, consistent with the previously published Bayesian results. The lowest mean FAP value was 0.04 at a frequency of 51.5 nHz; accounting for the trials factor gives $p \sim 1$.

We have shown that the $\mathcal{F}_p$ -statistic is a robust tool for detecting CWs in the presence of pulsar red noise and a GWB. We get similar results whether we marginalize over noise posteriors or use the maximum-likelihood values. This implies that the MLFP is sufficient enough to

calculate the significance of CWs. This is in contrast to frequentist searches for the GWB using the optimal statistic [30], where the common process is highly covariant with intrinsic pulsar red noise. It is possible that with more advanced red noise models, we could introduce a covariance between the CW and the intrinsic red noise. For future work, one can incorporate more flexible intrinsic red noise models. Since certain unmodeled noise processes can be mistaken for CWs, it will also be important to test the effect of newer noise models, such as chromatic noise [45–48] that may otherwise be mistaken for CWs. This would help remove biases from the $\mathcal{F}_p$ -statistic.

Frequentist searches with the $\mathcal{F}_p$ -statistic complement Bayesian searches. The $\mathcal{F}_p$ -statistic uses a simplified signal model, but is significantly faster to compute compared to performing a Bayesian analysis. With FASTFP the calculations can be accomplished in a matter of minutes using GPU acceleration, making it a useful tool for performing rapid searches for CWs across the PTA spectrum.

ACKNOWLEDGMENTS

We would like to thank Rand Burnette, Bjorn Larson, and Chiara Mingarelli for their comments and helpful insights. S.C.S. and S.J.V. are supported by NSF Award No. 2309246. G.E.F. is supported by National Aeronautics and Space Administration (NASA) Future Investigators in NASA Earth and Space Science and Technology Grant No. 80NSSC22K1591. C.A.W. acknowledges support from CIERA, the Adler Planetarium, and the Brinson Foundation through a CIERA-Adler postdoctoral fellowship. The authors are members of the NANOGrav Collaboration, which is supported by NSF Physics Frontiers Center Award No. 2020265.

DATA AVAILABILITY

The data that support the findings of this article are openly available [17,43].

APPENDIX A: INJECTED RED NOISE PARAMETERS

In Table I we present the intrinsic red noise parameters injected in our datasets.

APPENDIX B: CW PARAMETERS

In Table II we show the parameters governing the CW injections in both cases. The sky position is the most sensitive at 6 nHz from the NANOGrav 15 yr dataset [18].

TABLE I. The injected intrinsic red noise parameters (spectral index and log amplitude) for all the simulations, as well as the common process.

Pulsar	γ	$\log_{10} A$	Pulsar	γ	$\log_{10} A$
B1855 + 09	4.6094	-14.1708	J1730 - 2304	2.108	-18.859
B1937 + 21	4.0926	-13.633	J1738 + 0333	4.026	-14.210
B1953 + 29	1.894	-12.907	J1741 + 1351	3.629	-19.695
J0023 + 0923	0.3826	-14.182	J1744 - 1134	5.677	-18.795
J0030 + 0451	5.488	-15.012	J1745 + 1017	2.445	-11.921
J0340 + 4130	3.323	-15.860	J1747 - 4036	3.033	-12.791
J0406 + 3039	3.769	-16.043	J1751 - 2857	2.111	-18.693
J0437 - 4715	0.9381	-18.160	J1802 - 2124	0.604	-12.123
J0509 + 0856	0.231	-12.963	J1811 - 2405	2.093	-19.812
J0557 + 1551	6.183	-19.043	J1832 - 0836	0.587	-15.831
J0605 + 3757	0.075	-17.611	J1843 - 1113	2.308	-16.670
J0610 - 2100	2.549	-12.688	J1853 + 1303	0.428	-13.270
J0613 - 0200	2.977	-13.667	J1903 + 0327	1.376	-12.162
J0636 + 5128	3.911	-15.842	J1909 - 3744	2.215	-17.9584
J0645 + 5158	1.677	-13.538	J1910 + 1256	0.689	-15.539
J0709 + 0458	2.537	-12.857	J1911 + 1347	4.680	-15.506
J0740 + 6620	0.030	-14.636	J1918 - 0642	3.885	-17.894
J0931 - 1902	3.444	-15.315	J1923 + 2515	0.996	-18.762
J1012 + 5307	0.498	-12.565	J1944 + 0907	1.074	-13.066
J1012 - 4235	2.528	-13.671	J1946 + 3417	0.537	-12.441
J1022 + 1001	1.015	-12.106	J2010 - 1323	2.325	-15.567
J1024 - 0719	1.134	-15.699	J2017 + 0603	3.549	-18.739
J1125 + 7819	5.193	-17.020	J2033 + 1734	5.221	-15.429
J1312 + 0051	0.694	-12.776	J2043 + 1711	1.488	-17.604
J1453 + 1902	2.508	-19.971	J2124 - 3358	1.308	-19.499
J1455 - 3330	3.688	-16.338	J2145 - 0750	0.264	-12.814
J1600 - 3053	3.075	-15.765	J2214 + 3000	4.795	-16.653
J1614 - 2230	6.615	-18.599	J2229 + 2643	6.826	-16.640
J1630 + 3734	3.665	-14.084	J2234 + 0611	6.628	-15.630
J1640 + 2224	4.970	-16.783	J2234 + 0944	1.476	-15.908
J1643 - 1224	1.240	-12.296	J2302 + 4442	1.403	-15.363
J1705 - 1903	0.463	-12.073	J2317 + 1439	6.896	-17.790
J1713 + 0747	4.139	-16.175	J2322 + 2057	2.020	-14.796
J1719 - 1438	4.186	-16.637	CURN	4.333	-14.699

TABLE II. The injected parameters for the deterministic individual source.

Parameters	Injected value
$\mathcal{M}_c$	$10^9 M_\odot$
f_{cw}	8, 16 nHz
d_L	84.3 Mpc, 82 Mpc
θ	0
ϕ	$7\pi/4$
Φ_0	0
ι	0
ψ	0
SNR	5.36, 9.84

APPENDIX C: $\mathcal{F}_p$ RECOVERY

We show the recovery of the $\mathcal{F}_p$ values for each frequency across all simulations using $p - p$ diagrams in Fig. 11. We calculate the percentage of the recovered MLFP and recovered mean NMFP values with respect to the expected chi-squared distribution for all each of the 50 simulations. We then plot the cumulative fraction of simulations that recover percentages below a certain value [$P(\mathcal{F}_p^{\text{rec}} < \mathcal{F}_p)$].

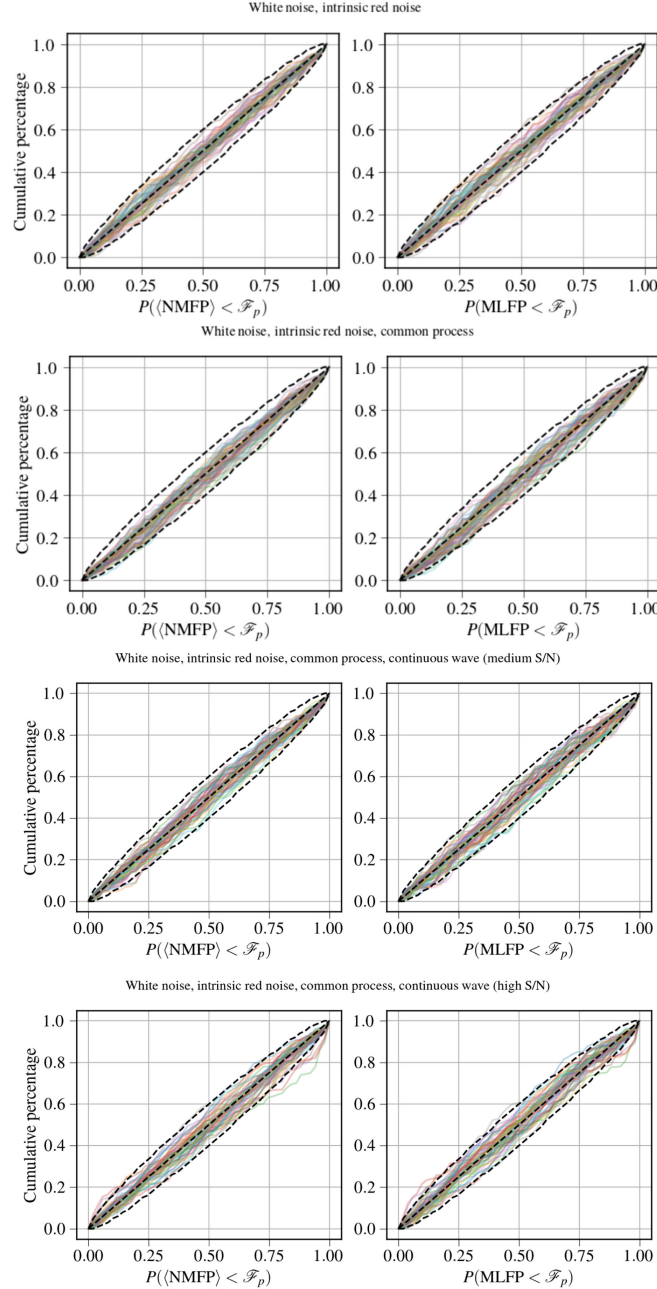


FIG. 11. $p - p$ diagrams for the four different types of simulations. The left panels show the recovery of the mean NMFP values, and the right panels show the recovery of the MLFP values. The horizontal axis is the recovered percentages and the vertical axis the cumulative fraction of simulations that recover that particular percentage. The dashed lines indicate a 95% confidence interval.

- [1] M. V. Sazhin, *Sov. Astron.* **22**, 36 (1978).
- [2] S. Detweiler, *Astrophys. J.* **234**, 1100 (1979).
- [3] R. S. Foster and D. C. Backer, *Astrophys. J.* **361**, 300 (1990).
- [4] D. N. Matsakis, J. H. Taylor, and T. M. Eubanks, *Astron. Astrophys.* **326**, 924 (1997).
- [5] J. P. W. Verbiest, M. Bailes, W. A. Coles, G. B. Hobbs, W. van Straten, D. J. Champion, F. A. Jenet, R. N. Manchester, N. D. R. Bhat, J. M. Sarkissian *et al.*, *Mon. Not. R. Astron. Soc.* **400**, 951 (2009).
- [6] M. Rajagopal and R. W. Romani, *Astrophys. J.* **446**, 543 (1995).
- [7] A. H. Jaffe and D. C. Backer, *Astrophys. J.* **583**, 616 (2003).
- [8] J. S. B. Wyithe and A. Loeb, *Astrophys. J.* **590**, 691 (2003).
- [9] P. A. Rosado, A. Sesana, and J. Gair, *Mon. Not. R. Astron. Soc.* **451**, 2417 (2015).
- [10] G. Agazie, A. Anumalapudi, A. M. Archibald, Z. Arzoumanian, P. T. Baker, B. Bécsy, L. Blecha, A. Brazier, P. R. Brook, S. Burke-Spolaor *et al.*, *Astrophys. J. Lett.* **951**, L8 (2023).
- [11] J. Antoniadis, P. Arumugam, S. Arumugam, S. Babak, M. Bagchi, A. S. Bak Nielsen, C. G. Bassa, A. Bathula *et al.* (EPTA, InPTA Collaborations), *Astron. Astrophys.* **678**, A50 (2023).
- [12] D. J. Reardon, A. Zic, R. M. Shannon, G. B. Hobbs, M. Bailes, V. Di Marco, A. Kapur, A. F. Rogers, E. Thrane, J. Askew *et al.*, *Astrophys. J. Lett.* **951**, L6 (2023).
- [13] H. Xu, S. Chen, Y. Guo, J. Jiang, B. Wang, J. Xu, Z. Xue, R. Nicolas Caballero, J. Yuan, Y. Xu *et al.*, *Res. Astron. Astrophys.* **23**, 075024 (2023).
- [14] M. T. Miles, R. M. Shannon, D. J. Reardon, M. Bailes, D. J. Champion, M. Geyer, P. Gitika, K. Grunthal, M. J. Keith, M. Kramer *et al.*, *Mon. Not. R. Astron. Soc.* **536**, 1489 (2025).
- [15] B. Bécsy, N. J. Cornish, and L. Z. Kelley, *Astrophys. J.* **941**, 119 (2022).
- [16] K. Aggarwal, Z. Arzoumanian, P. T. Baker, A. Brazier, M. R. Brinson, P. R. Brook, S. Burke-Spolaor, S. Chatterjee, J. M. Cordes, N. J. Cornish *et al.*, *Astrophys. J.* **880**, 116 (2019).
- [17] Z. Arzoumanian, P. T. Baker, L. Blecha, H. Blumer, A. Brazier, P. R. Brook, S. Burke-Spolaor, B. Bécsy, J. A. Casey-Clyde, M. Charisi *et al.*, *Astrophys. J. Lett.* **951**, L28 (2023).
- [18] G. Agazie, A. Anumalapudi, A. M. Archibald, Z. Arzoumanian, P. T. Baker, B. Bécsy, L. Blecha, A. Brazier, P. R. Brook, S. Burke-Spolaor *et al.*, *Astrophys. J. Lett.* **951**, L50 (2023).
- [19] E. C. Gardiner, B. Bécsy, L. Z. Kelley, and N. J. Cornish, *Astrophys. J.* **988**, 222 (2025).
- [20] P. Jaranowski, A. Królak, and B. F. Schutz, *Phys. Rev. D* **58**, 063001 (1998).
- [21] N. J. Cornish and E. K. Porter, *Classical Quantum Gravity* **24**, 5729 (2007).
- [22] A. Sesana and A. Vecchio, *Phys. Rev. D* **81**, 104008 (2010).
- [23] S. Babak and A. Sesana, *Phys. Rev. D* **85**, 044034 (2012).
- [24] J. A. Ellis, X. Siemens, and J. D. E. Creighton, *Astrophys. J.* **756**, 175 (2012).
- [25] Z. Arzoumanian, A. Brazier, S. Burke-Spolaor, S. J. Chamberlin, S. Chatterjee, J. M. Cordes, P. B. Demorest, X. Deng, T. Dolch, J. A. Ellis *et al.*, *Astrophys. J.* **794**, 141 (2014).
- [26] J. Antoniadis, P. Arumugam, S. Arumugam, S. Babak, M. Bagchi, A. S. Bak Nielsen, C. G. Bassa, A. Bathula *et al.* (EPTA, InPTA Collaborations), *Astron. Astrophys.* **690**, A118 (2024).
- [27] K. Furusawa, S. Kuroyanagi, and K. Ichiki, *Mon. Not. R. Astron. Soc.* **543**, 1010 (2025).
- [28] K. Aggarwal, Z. Arzoumanian, P. T. Baker, A. Brazier, M. R. Brinson, P. R. Brook, S. Burke-Spolaor, S. Chatterjee, J. M. Cordes, N. J. Cornish *et al.*, *Astrophys. J.* **880**, 116 (2019).
- [29] 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).
- [30] S. J. Vigeland, K. Islo, S. R. Taylor, and J. A. Ellis, *Phys. Rev. D* **98**, 044003 (2018).
- [31] M. F. Alam, Z. Arzoumanian, P. T. Baker, H. Blumer, K. E. Bohler, A. Brazier, P. R. Brook, S. Burke-Spolaor, K. Caballero, R. S. Camuccio *et al.*, *Astrophys. J. Suppl. Ser.* **252**, 4 (2021).
- [32] R. S. Lynch, *J. Phys. Conf. Ser.* **610**, 012017 (2015).
- [33] H. Wahlquist, *Gen. Relativ. Gravit.* **19**, 1101 (1987).
- [34] K. J. Lee, N. Wex, M. Kramer, B. W. Stappers, C. G. Bassa, G. H. Janssen, R. Karuppusamy, and R. Smits, *Mon. Not. R. Astron. Soc.* **414**, 3251 (2011).
- [35] V. Corbin and N. J. Cornish, *arXiv:1008.1782*.
- [36] R. W. Hellings and G. S. Downs, *Astrophys. J. Lett.* **265**, L39 (1983).
- [37] S. R. Taylor, E. A. Huerta, J. R. Gair, and S. T. McWilliams, *Astrophys. J.* **817**, 70 (2016).
- [38] P. Jaranowski and A. Królak, *Phys. Rev. D* **61**, 062001 (2000).
- [39] J. Bradbury, R. Frostig, P. Hawkins, M. J. Johnson, C. Leary, D. Maclaurin, G. Necula, A. Paszke, J. VanderPlas, S. Wanderman-Milne *et al.*, JAX: Composable transformations of python + numpy programs (2018), URL <http://github.com/google/jax>.
- [40] J. A. Ellis, M. Vallisneri, S. R. Taylor, and P. T. Baker, ENTERPRISE: Enhanced Numerical Toolbox Enabling a Robust Pulsar Inference Suite, Zenodo (2020).
- [41] S. R. Taylor, P. T. Baker, J. S. Hazboun, J. Simon, and S. J. Vigeland, *enterprise_extensions* (2021), v2.3.3, URL https://github.com/nanograv/enterprise_extensions.
- [42] J. Ellis and R. van Haasteren, *jellis18/ptmcmcsampler*: Official release (2017), URL [10.5281/zenodo.1037579](https://doi.org/10.5281/zenodo.1037579).
- [43] G. Agazie, M. F. Alam, A. Anumalapudi, A. M. Archibald, Z. Arzoumanian, P. T. Baker, L. Blecha, V. Bonidie, A. Brazier, P. R. Brook *et al.*, *Astrophys. J. Lett.* **951**, L9 (2023).
- [44] M. Vallisneri, P. M. Meyers, K. Chatziioannou, and A. J. K. Chua, *Phys. Rev. D* **108**, 123007 (2023).
- [45] J. S. Hazboun, J. Simon, J. Baier, B. Larsen, D. J. Oliver, P. T. Baker, B. Bécsy, S. Chen, A. Diaz Hernandez, J. A. Ellis *et al.*, *arXiv:2511.22597*.

- [46] L. Lentati, R. M. Shannon, W. A. Coles, J. P. W. Verbiest, R. van Haasteren, J. A. Ellis, R. N. Caballero, R. N. Manchester, Z. Arzoumanian, S. Babak *et al.*, [Mon. Not. R. Astron. Soc. **458**, 2161 \(2016\)](#).
- [47] D. J. Reardon, A. Zic, R. M. Shannon, V. Di Marco, G. B. Hobbs, A. Kapur, M. E. Lower, R. Mandow, H. Middleton, M. T. Miles *et al.*, [Astrophys. J. Lett. **951**, L7 \(2023\)](#).
- [48] B. Larsen, C. M. F. Mingarelli, J. S. Hazboun, A. Chalumeau, D. C. Good, J. Simon, G. Agazie, A. Anumalapudi, A. M. Archibald, Z. Arzoumanian *et al.*, [Astrophys. J. **972**, 49 \(2024\)](#).