

Searching for anisotropic gravitational-wave backgrounds using pulsar timing arrays

Stephen R. Taylor^{*} and Jonathan R. Gair[†]

Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA, United Kingdom

(Received 19 July 2013; published 1 October 2013)

We present the results of simulated injections testing the first Bayesian search pipeline capable of investigating the angular structure of a gravitational-wave (GW) background influencing pulsar signals. A stochastic background of GWs from the incoherent superposition of many inspiraling supermassive black hole binaries at nHz frequencies is likely to be the dominant GW signal detectable by pulsar timing arrays (PTAs). Even though one might expect a background composed of a high-redshift cosmological population of sources to be fairly isotropic, deviations from isotropy may be indicative of local GW hotspots or some form of continuous anisotropy in the angular distribution of GW power. A gravitational-wave background (GWB) induces time-of-arrival deviations in pulsar signals which are correlated between separated pulsars. In an isotropic background this cross correlation follows a distinctive relationship, known as the Hellings and Downs curve, that depends only on the angular separation of the pulsars. If the background is anisotropic, the cross correlation is different, but predictable, and also depends on the absolute position of the pulsars. By simulating data sets containing GWBs with various anisotropic configurations, we have explored the prospects for constraining anisotropy using near future data. We find that at moderate to high signal to noise ratio the assumption of isotropy is no longer an appropriate description of the simulated background. Furthermore, we can recover the nature of the injected anisotropy in a Bayesian parameter-estimation search, and propose a prior on the anisotropy search space motivated by the physicality of the implied distribution of sources. The TEMPO2 plugins used to generate pulsar time-of-arrival data sets influenced by anisotropic GWBs can be freely downloaded from the *tempo2-anisotropy-plugins* GitHub repository.

DOI: 10.1103/PhysRevD.88.084001

PACS numbers: 04.30.-w, 04.80.Nn, 97.60.Gb, 95.55.Ym

I. INTRODUCTION

The next several years should see the true beginnings of observational gravitational-wave (GW) astronomy. The ground-based interferometers, such as aLIGO [1], AdVirgo [2], and KAGRA [3], are expected to be operating at full design sensitivity by the end of this decade, although the first science runs may occur several years before that [4]. These detectors will be sensitive to the chirping signal of stellar mass compact-binary inspirals in the frequency range $\sim 10\text{--}10^3$ Hz. Additionally, it is hoped that a space-based interferometer with arm lengths of $\sim 10^9$ m, such as eLISA/NGO [5,6], will be operable by the end of the 2020s, and sensitive to GWs in the frequency range $\sim 0.1\text{--}100$ mHz generated by binaries of massive black holes (MBHs) and the extreme-mass-ratio inspirals of stellar mass compact objects into MBHs.

Gravitational waves can also be detected through their effect on pulsar timing residuals. Several recent studies have claimed that pulsar timing arrays (PTAs) [7] are on the cusp of detecting a nHz stochastic GW background (GWB) [8–10]. PTAs exploit the timing accuracy of millisecond pulsars [11,12] to measure the influence of a GW perturbing the space-time metric along each pulsar-Earth line of sight [13–16]. The frequency range that a PTA is sensitive to is set by the pulsar observational time span ($f_{\text{low}} \sim 1/T$), and

the regularity of pulsar observations [$f_{\text{high}} \sim 1/(2\Delta T)$]. The pulsar array essentially forms a kpc GW detector which is complementary to other GW experiments.

In practice we must disentangle the influence of a GW on the time of arrival (TOA) of pulses from other stochastic and deterministic influences. These include the quadratic spindown of the pulsar itself, as its rotational energy is extracted to power the EM outflow. There are also a variety of noise sources, such as clock noise, receiver noise and dispersion measure effects. The deterministic parameters can be fit out of the raw TOAs to leave the timing residuals, which include all unmodeled phenomena, such as GWs.

The dominant GW signal expected in the PTA band will result from the incoherent superposition of signals from inspiraling Supermassive Black-Hole Binaries (SMBHBs), creating a stochastic background. Essentially, the GW emission from many systems piles up in the same frequency bin such that, at low frequency at least, no single source can be individually resolved.. At nHz frequencies this creates a characteristic strain spectrum that is approximately a power law [17–20]. However, at higher frequencies ($10^{-8}\text{--}10^{-7}$ Hz), the signal may become dominated by several close, bright binaries, such that the stochasticity of the signal breaks down [21] and the shape of the spectrum will show a stronger dependence on the particular MBH assembly scenario in our Universe.

The characteristic strain spectrum of a GWB is defined in terms of the fractional contribution of the background to the energy density of the Universe,

^{*}staylor@ast.cam.ac.uk

[†]jgair@ast.cam.ac.uk

$$\Omega_{\text{GW}}(f) = \frac{1}{\rho_c} \frac{d\rho_{\text{GW}}(f)}{d(\ln f)} = \frac{\pi}{4} \frac{f^2 h_c(f)^2}{\rho_c}, \quad (1)$$

where f is the observed GW frequency, $\rho_c = 3H^2/8\pi G$ is the energy density required for a flat Universe, and $h_c(f)$ is the characteristic strain of the GW background in a frequency interval centered at f .

The characteristic strain spectrum of a GW background resulting from inspiraling SMBHBs is approximately $h_c(f) \propto f^{-2/3}$, and to a good approximation the strain spectrum of a background from other sources can also be described by a power law. Some measurable primordial background contributions may have a power-law index of -1 [22–25], while the background from decaying cosmic strings [26–29] may have $-7/6$ [30]. For most models of interest, we can describe the strain spectrum of a stochastic GW background by $h_c(f) = A(f/\text{yr}^{-1})^\alpha$ [31]. This characteristic strain spectrum is related to the one-sided power spectral density of the GWB-induced TOA deviations by

$$S(f) = \frac{1}{12\pi^2} \frac{1}{f^3} h_c(f)^2 = \frac{A^2}{12\pi^2} \left(\frac{f}{\text{yr}^{-1}} \right)^{-\gamma} \text{yr}^3, \quad (2)$$

where $\gamma \equiv 3 - 2\alpha$.

If a GWB is produced by processes occurring in the early Universe, or a superposition of high-redshift signals, then one would expect the resulting background to be reasonably isotropic. However, if the background signal is dominated by close, bright sources over the entire PTA-sensitivity band (as studied by Ravi *et al.* [32]), or even just at high frequencies, the breakdown of stochasticity and resulting deviation from isotropy motivates a more general characterization of the angular distribution of GW power on the sky.

Over the last several years, constraints on the amplitude of an *isotropic* GWB have been published by the three major PTA collaborations [33–35], the European Pulsar Timing Array (EPTA) [36], the North American Nanohertz Observatory for Gravitational Waves (NANOGrav) [37], and the Parkes Pulsar Timing Array (PPTA) [38]. The International Pulsar Timing Array (IPTA) [39] consortium combines these three efforts. The techniques used by these groups rely on the fact that a GWB induces correlated TOA deviations between separated pulsars, which provides leverage against the undesirable noise processes. In fact, for an isotropic GWB this cross correlation has a unique, distinctive angular signature dependant only on the angular separation of the pulsars, and is commonly referred to as the Hellings and Downs curve [40].

Recently, Cornish and Sesana [41] proposed a cross-correlation statistic aimed at targeting the brightest sources in each PTA frequency band to account for anisotropy in the signal power. This statistic modifies the existing isotropic correlation techniques to account for finiteness of the source population. Essentially, instead of integrating the directional correlation function over the sky to produce

the Hellings and Downs curve, this modified statistic sums the cross-correlation function over the brightest sources to infer their sky positions.

Alternatively, Mingarelli *et al.* [42] have generalized the existing isotropic cross-correlation technique to account for arbitrary levels of anisotropy in a GWB. This is achieved by decomposing the angular power distribution in terms of spherical harmonics. Isotropy is represented by the monopole solution. For higher multipoles, the cross correlation is no longer a simple function of pulsar angular separation, but rather depends on the position of each pulsar relative to orientation of the background anisotropy. Thus the concept of a one-parameter relationship to describe correlations induced by a GWB is no longer appropriate, and the relationship will instead become PTA- and GWB-specific, depending on the observed pulsars and the amount of anisotropy in the GWB.

In this paper we develop and test the first anisotropic Bayesian search pipeline for the characterization of a GWB using PTAs, employing the previously mentioned generalized spherical-harmonic decomposition of the cross correlation. We have developed a suite of new plugins for the popular pulsar timing analysis package TEMPO2 [43–45], which can be used to inject GWBs with any user-defined anisotropy. For this first investigation of measuring the anisotropic correlations, we inject backgrounds composed of typically 10^4 sources, with sources placed according to various tested angular distributions. As such, we are investigating continuous anisotropy in the distribution of source positions, rather than anisotropy through finiteness of the background, which we defer to future work.

This paper is organized as follows. In Sec. II we perform some theoretical estimates of the expected level of anisotropy in a background composed of a superposition of many GW signals, where the anisotropy either derives from varying source brightness or sparseness of the source population. In Sec. III we review the theory of time and angular correlations of GWB-induced timing residuals in pulsar signals, and outline a generalized correlation formalism for modeling correlations in anisotropic backgrounds. The theory of Bayesian inference is briefly reviewed in Sec. IV, followed by a description in Sec. V of the data-analysis stages involved in the search for a stochastic GWB in pulsar data sets. A pipeline for producing simulated pulsar data sets containing an anisotropic GWB signal is described in Sec. VI, followed by the results of Bayesian parameter estimation and evidence evaluation on these simulated data sets in Sec. VII, with our conclusions in Sec. VIII.

II. THE EXPECTED LEVEL OF ANISOTROPY

We wish to motivate a Bayesian search pipeline which is generalized to arbitrary levels of anisotropy in the power distribution of the GW background. We take a realistic population of SMBHBs, generated by grafting baryonic physics onto the dark-matter haloes of a Millennium

Simulation realization, and extract those systems with observed GW frequencies which fall within PTA observation frequencies [10]. These are typically massive and close systems ($z < 2$). A typical catalog contains $\sim 2 \times 10^4$ sources, each with associated chirp masses, \mathcal{M} ; observed GW frequencies, f ; and redshifts, z . Most importantly, the source positions are distributed isotropically on the sky. Hence, with this many sources making up our background we would expect any anisotropy to derive from varying source brightnesses. The radiated power in GWs of each binary in the source rest frame is estimated using the quadrupole formula,

$$\frac{dE}{dt} = \frac{32}{5} [\pi \mathcal{M} f(1+z)]^{10/3}, \quad (3)$$

where $G = c = 1$. The radiated energy flux in the observer frame is then given by

$$\mathcal{F} = \frac{dE}{dt} \times \frac{1}{4\pi D_L(z)^2}, \quad (4)$$

where D_L is the source luminosity distance. The energy density in GWs, ρ , received from each source is then simply given by \mathcal{F}/c .

We now perform a spherical-harmonic decomposition of the energy density from a GW population, for the purposes of evaluating the angular power spectrum. Thus,

$$\rho(\hat{\Omega}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l c_{lm} Y_{lm}(\hat{\Omega}) \quad (5)$$

from which we compute the anisotropy coefficients,

$$c_{lm} = \int_{S^2} d\hat{\Omega} \rho(\hat{\Omega}) Y_{lm}(\hat{\Omega}). \quad (6)$$

We define the observed angular power spectrum as $C_l = \sum_m |c_{lm}|^2 / (2l + 1)$. Since we are dealing with a population of point sources, the energy-density distribution is just a sequence of delta functions at the source positions, such that the spherical-harmonic decomposition becomes

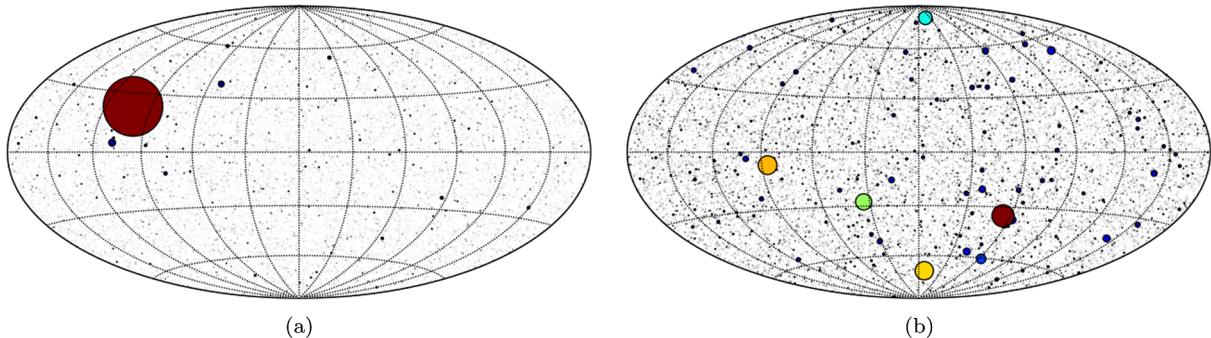


FIG. 1 (color online). Skymaps of GW source populations generated by evolving a population of SMBHB systems. There are $\sim 2 \times 10^4$ systems in each catalog, which are typically massive (10^7 – $10^{10} M_\odot$) and close ($z < 2$). The relative size and color of points within each skymap is indicative of the GW energy flux from each system. The GW signal from the first data set in (a) is clearly dominated by one very bright source. In the second data set (b) we have several bright sources, however no outliers as in the first data set.

$$c_{lm} = \sum_{i=1}^N \rho_i Y_{lm}(\hat{\Omega}_i) \quad (7)$$

which is just a summation over the spherical-harmonic functions evaluated at the N source positions, and weighted by the GW energy-density of each source.

We analyze two catalogs, whose skymaps are shown in Fig. 1. The relative size and color of each source is indicative of the GW energy density, where larger and redder denotes a brighter source. The first data set clearly includes a source which swamps the signal, while the second data set is dominated by several bright sources. Figure 2(a) shows the angular power spectrum of the first data set, evaluated up to $l = 50$. We see that the single brightest source swamps the power at all l , such that we are almost trivially satisfying Unsöld's Theorem by having $[C_l/C_0] = 1$ at each l . If we remove the brightest source, then we get a power spectrum in which the power is predominantly in the monopole (isotropic), with typically less than 10% of the isotropic power in the other l values. This small level of power in other l values most likely derives from the few next-brightest sources, since if we fix all source brightness to the same value we obtain a power fraction of $\sim 10^{-5}$ in $l \neq 0$ values, which is what we expect when analyzing a large population of equal-brightness sources which have been isotropically distributed on the sky.

The second data set has several very bright sources, but no outliers like in the first data set. When analyzed, most of the power lies in the monopole, with typically less than 10% of the isotropic power in modes with $l \neq 0$. However, if we simulate source sparseness, or some form of intrinsic anisotropy through source clustering, then the potential level of anisotropy is much higher. In Fig. 2(b) the data set has been shuffled and split into subpopulations. We compute the angular power spectrum of each subpopulation, and display the 95% upper limit of the power in each l derived from the ensemble of subpopulations. With a very sparse population the power in the modes with $l \neq 0$ could conceivably be $\leq 80\%$ of the isotropic power. This anisotropic power

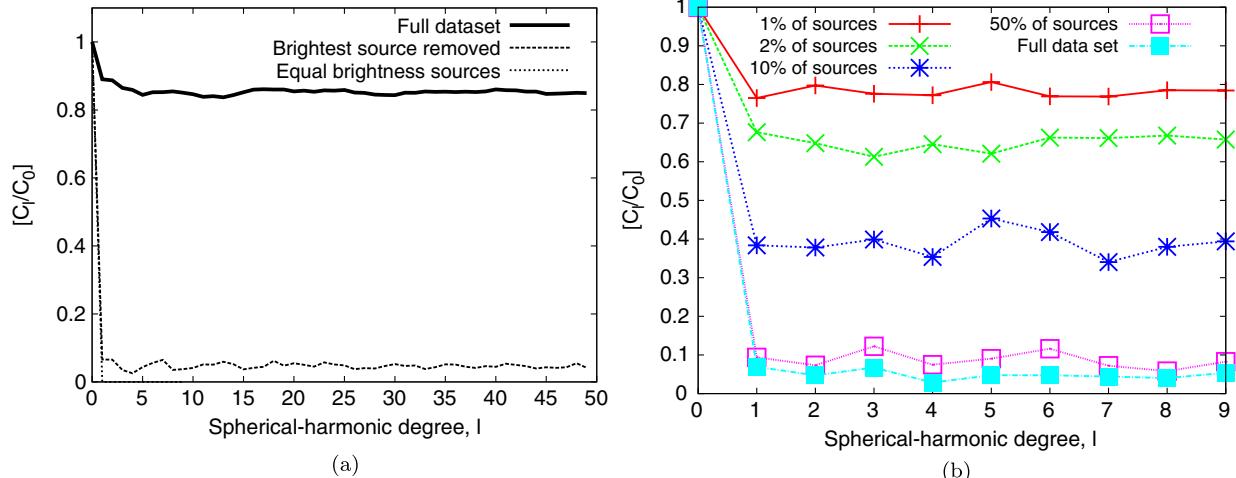


FIG. 2 (color online). We perform a spherical-harmonic decomposition of the distribution of the GW energy density on the sky resulting from a population of inspiraling SMBHB systems, and compute the angular power spectrum, C_l . In (a) the influence of one very bright source in skymap (a) of Fig. 1 clearly swamps all other influences, such that we almost trivially satisfy Unsöld's theorem. If we remove the brightest source, then the anisotropic power drops to less than 0.1 relative to the monopole (isotropic). Even this level of anisotropy is most likely due to the second-brightest sources, since if we assume equal brightness sources the $l \neq 0$ power drops to $\sim 10^{-5}$. This is expected for a large population of isotropically distributed source positions. In (b) we analyze skymap (b) of Fig. 1. The full data set is predominantly isotropic; however if we simulate source sparseness or intrinsic anisotropy by analyzing subpopulations, then the level of anisotropy can be quite high.

fraction tends to the value for the full data set as we include more sources in the subpopulations.

Thus, if we have several very bright sources dominating our background, or some form of intrinsic anisotropy through source clustering, then the level of anisotropy in the sky distribution of GW energy density could conceivably be high. In the following, we simulate anisotropic backgrounds through a continuous anisotropy in GW source positions, and test the robustness of a Bayesian anisotropic search pipeline.

III. CORRELATIONS INDUCED BY A GWB

In this section we provide a brief overview of the time and angular correlations in the GW-induced deviations to the TOA of pulsar signals. We reiterate some of the formalism of Mingarelli *et al.* [42] for ease of reference.

A. Cross correlating the timing residuals

The redshift of a signal from a pulsar in the direction of unit vector \hat{p} , induced by the passage of a *single* GW propagating in the direction of $\hat{\Omega}$ is [46,47]

$$z(t, \hat{\Omega}) = \frac{1}{2} \frac{\hat{p}^a \hat{p}^b}{1 + \hat{\Omega} \cdot \hat{p}} \Delta h_{ab}(t, \hat{\Omega}), \quad (8)$$

where $\Delta h_{ab} \equiv h_{ab}(t_e, \hat{\Omega}) - h_{ab}(t_p, \hat{\Omega})$, is the difference in the metric perturbation at the solar system barycenter, $h_{ab}(t_e, \hat{\Omega})$, and at the pulsar, $h_{ab}(t_p, \hat{\Omega})$.

This frequency shift is integrated over time to give the induced timing residuals, which describe the perturbation to the TOA of pulses from a given pulsar,

$$r(t) \equiv \int_0^t z(t') dt'. \quad (9)$$

In the transverse-traceless (TT) gauge the perturbation of the metric caused by a mass quadrupole moment can be described in terms of the linear superposition of the amplitudes of the two GW polarizations permitted in general relativity (GR). Hence, the GR metric perturbation can be written as

$$h_{ab}(t, \hat{\Omega}) = h_+(t, \hat{\Omega}) e_{ab}^+(t, \hat{\Omega}) + h_\times(t, \hat{\Omega}) e_{ab}^\times(t, \hat{\Omega}) \quad (10)$$

where $e_{ab}^{+,\times}(\hat{\Omega})$ are the polarization tensors, defined in terms of orthonormal basis vectors around $\hat{\Omega}$,

$$e_{ab}^+(\hat{\Omega}) = \hat{m}_a \hat{m}_b - \hat{n}_a \hat{n}_b \quad e_{ab}^\times(\hat{\Omega}) = \hat{m}_a \hat{n}_b + \hat{n}_a \hat{m}_b. \quad (11)$$

In a spherical-polar coordinate system, where the direction of GW propagation is $\hat{\Omega} = [\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta]$, the polarization basis vectors take the following form:

$$\begin{aligned} \hat{m} &= [\sin \phi, -\cos \phi, 0] \\ \hat{n} &= [\cos \theta \cos \phi, \cos \theta \sin \phi, -\sin \theta]. \end{aligned} \quad (12)$$

For the purposes of the following, we can also express the GW metric perturbation at a position in our coordinate system, \vec{x} , in terms of a plane-wave expansion. In a coordinate system centered at the solar system barycenter, with the pulsar some distance L away (such that $t_p = t_e - L = t - L$, $\hat{x}_e = 0$, and $\hat{x}_p = L \hat{p}$) this has the form

$$h_{ab}(t, \vec{x}) \equiv \sum_{A=+, \times} \int_{-\infty}^{\infty} df \int_{S^2} d\hat{\Omega} h_A(f, \hat{\Omega}) e^{2\pi if(t-\hat{\Omega} \cdot \vec{x})} e_{ab}^A(\hat{\Omega}). \quad (13)$$

Using this, and integrating the $\Delta h_{ab}(t, \hat{\Omega})$ term in Eq. (8) over all positions of sources constituting the stochastic GWB, gives

$$\Delta h_{ab}(t) = \sum_{A=+, \times} \int_{-\infty}^{\infty} df \int_{S^2} d\hat{\Omega} e_{ab}^A(\hat{\Omega}) h_A(f, \hat{\Omega}) e^{2\pi ift} \times [1 - e^{-2\pi ifL(1+\hat{\Omega} \cdot \hat{p})}] \quad (14)$$

where \hat{p} and L are the position and distance to the pulsar, respectively. The redshift of a signal induced by a stochastic GWB can now be written as

$$z(t) = \sum_{A=+, \times} \int_{-\infty}^{\infty} df \int_{S^2} d\hat{\Omega} F^A(\hat{\Omega}) h_A(f, \hat{\Omega}) e^{2\pi ift} \times [1 - e^{-2\pi ifL(1+\hat{\Omega} \cdot \hat{p})}] \quad (15)$$

where the antenna beam pattern for each polarization in the PTA formalism is

$$F^A(\hat{\Omega}) \equiv \frac{1}{2} \frac{\hat{p}^a \hat{p}^b}{1 + \hat{\Omega} \cdot \hat{p}} e_{ab}^A(\hat{\Omega}). \quad (16)$$

GW-detection pipelines used in PTAs search for correlated perturbations induced by GWs affecting different pulsars. This influence is manifested through the perturbation to the expected TOA of a given pulse caused by the passage of the GWs, such that what we actually measure are a set of TOA residuals. These residuals encode information about all unmodeled phenomena, such as detector noise, intrinsic pulsar spin noise and pulse phase jitter. Crucially though, these sorts of influences are uncorrelated between different pulsars. The residuals induced by GWs in different pulsars will be correlated, which provides us the leverage required to separate these from noise processes. We will briefly discuss the details of how raw TOAs are processed into TOA residuals in Sec. V.

Explicitly evaluating Eq. (9) we find

$$r(t) = \frac{i}{2\pi} \sum_{A=+, \times} \int_{-\infty}^{\infty} df \int_{S^2} d\hat{\Omega} F^A(\hat{\Omega}) \frac{h_A(f, \hat{\Omega})}{f} \times [1 - e^{-2\pi ifL(1+\hat{\Omega} \cdot \hat{p})}] [1 - e^{2\pi ift}] \quad (17)$$

To determine the correlation function between the residual at time t_0 in pulsar a with the residual at time $(t_0 + t)$ in pulsar b , $\langle r_a^*(t_0) r_b(t_0 + t) \rangle$, we replace the statistical ensemble average with a time average, where the maximum time over which we can average is the light-travel time to the pulsar. We remove the unobservable constant offset from $r(t)$ [48]. Furthermore, if we have a background of GWs composed of a superposition of individually unresolvable single sources, and such a background is stationary, Gaussian and independently polarized, we have [49, 50]

$$\langle h_A^*(f, \hat{\Omega}) h_{A'}(f', \hat{\Omega}') \rangle = H(f) P(\hat{\Omega}) \delta_{AA'} \delta(f - f') \delta^2(\hat{\Omega}, \hat{\Omega}') \quad (18)$$

where $H(f)$ describes the spectral content of the background radiation. Also, $P(\hat{\Omega})$ is the angular distribution of GW power on the sky, which we interpret here as the angular distribution of sources comprising the background. Finally,

$$\langle r_a^*(t_0) r_b(t_0 + t) \rangle = \int_0^{\infty} df \Gamma_{ab}(f) \frac{H(f)}{f^2} e^{2\pi ift} \quad (19)$$

where

$$\begin{aligned} \Gamma_{ab}(f) &\equiv \int_{S^2} d\hat{\Omega} P(\hat{\Omega}) \kappa_{ab}(f, \hat{\Omega}) \left[\sum_{A=+, \times} F_a^A(\hat{\Omega}) F_b^A(\hat{\Omega}) \right] \\ \kappa_{ab}(f, \hat{\Omega}) &\equiv [1 - e^{2\pi ifL_a(1+\hat{\Omega} \cdot \hat{p}_a)}] [1 - e^{-2\pi ifL_b(1+\hat{\Omega} \cdot \hat{p}_b)}]. \end{aligned} \quad (20)$$

In this case $\Gamma_{ab}(f)$ takes on the role of the “overlap-reduction function” (ORF) due to noncolocated, nonaligned detectors, which is often seen in the LIGO/LISA literature.

The one-sided power spectrum of the induced residuals is related to the autocorrelation function via the Wiener-Kinchin theorem,

$$S(f) = \int_0^{\infty} \langle r_a^*(t_0) r_b(t_0 + t) \rangle e^{-2\pi ift} dt. \quad (21)$$

B. Angular correlations induced by a GWB

We now shift attention to the cross correlation $\Gamma_{ab}(f)$ between pulsars a and b . If we have a GWB composed of a superposition of individually unresolvable single sources, then we can express their angular distribution on the sky as a decomposition in spherical harmonics. For the remainder of this analysis we decompose in terms of the real spherical harmonics, Y_{lm} , which are formed by the following linear combinations of their complex analogues, Y_l^m :

$$Y_{lm} = \begin{cases} \frac{1}{\sqrt{2}} [Y_l^m + (-1)^m Y_l^{-m}] & m > 0 \\ Y_l^0 & m = 0 \\ \frac{1}{i\sqrt{2}} [Y_l^{-m} - (-1)^m Y_l^m] & m < 0 \end{cases} \quad (22)$$

such that the angular distribution of the GWB radiation is expressed as

$$P(\hat{\Omega}) \equiv \sum_{l=0}^{\infty} \sum_{m=-l}^l c_{lm} Y_{lm}(\hat{\Omega}). \quad (23)$$

Examining the form of $\Gamma_{ab}(f)$ in Eq. (20), we see that the overlap reduction function can also be decomposed into a set of “correlation basis functions,”

$$\Gamma_{ab}(f) = \sum_{l=0}^{\infty} \sum_{m=-l}^l c_{lm} \Gamma_{lm}^{(ab)}(f) \quad (24)$$

where

$$\Gamma_{lm}^{(ab)}(f) \equiv \int_{S^2} d\hat{\Omega} Y_{lm}(\hat{\Omega}) \kappa_{ab}(f, \hat{\Omega}) \left[\sum_{A=+, \times} F_a^A(\hat{\Omega}) F_b^A(\hat{\Omega}) \right] \quad (25)$$

are the generalized ORFs, or correlation basis functions.

Using the formalism of Allen and Ottewill [51], we define a “cosmic frame” and a “computational frame.” In the former, pulsars have their usual right ascension (RA) and declination (DEC) positions on the sky. The latter frame places one pulsar along the z axis of the coordinate system, and the other pulsar in the xz -plane. This choice permits simple analytic expressions for the $\Gamma_{lm}^{(ab)}(f)$ to be determined, but we must remember that each distinct pair of pulsars defines a unique computational frame, so we must rotate the computed values of $\Gamma_{lm}^{(ab)}(f)$ back into the common cosmic frame.

We now quote the computational-frame correlation basis functions from Mingarelli *et al.* [42], which have been converted to their real form. Unless otherwise stated, we assume that $L_a = L_b$ and we are in the large fL limit, such that $\kappa_{ab}(f, \hat{\Omega}) \sim 1 + \delta_{ab}$ (where δ_{ab} is the Kronecker delta), and therefore $\Gamma_{ab}(f)$ is independent of f . Let the angular separation between pulsars a and b be ζ , and $\alpha = (1 + \cos \zeta)$ and $\beta = (1 - \cos \zeta)$.

$$\Gamma_{00} = N \frac{\sqrt{\pi}}{2} \left[1 + \frac{\cos \zeta}{3} + 4(1 - \cos \zeta) \ln \left(\sin \frac{\zeta}{2} \right) \right] \kappa_{ab}. \quad (26)$$

The monopole ($l = m = 0$) is the Hellings and Downs solution, as expected. We choose the normalization to be $N = 3/8\pi$ such that a pure isotropic background with $c_{00} = 2\sqrt{\pi}$ will have $\Gamma_{aa}(f) = c_{00}\Gamma_{00}^{(aa)}(f) = 1$.

In the following we refer to Γ_{lm} as $\Gamma_{l,m}$ to easily distinguish negative m values. The dipole correlation basis functions have the following form:

$$\begin{aligned} \Gamma_{1,-1} &= 0 \\ \Gamma_{1,0} &= -N \frac{1}{2} \sqrt{\frac{\pi}{3}} \left[\alpha + 3\beta \left(\alpha + 4 \ln \left(\sin \frac{\zeta}{2} \right) \right) \right] \kappa_{ab} \\ \Gamma_{1,1} &= N \frac{1}{2} \sqrt{\frac{\pi}{3}} \sin \zeta \ln \left[1 + 3\beta \left(1 + \frac{4}{\alpha} \ln \left(\sin \frac{\zeta}{2} \right) \right) \right] \kappa_{ab}. \end{aligned} \quad (27)$$

An alternative phrasing of the correlation induced by a dipole anisotropy in the GWB was derived in Anholm *et al.* [46], but as shown in Mingarelli *et al.* [42] it is completely consistent with the formalism above:

$$\begin{aligned} \Gamma_{ab,\text{dip}} &= N\pi(\cos \delta_a + \cos \delta_b) \\ &\times \left(\cos \zeta - \frac{4}{3} - 4\tan^2 \left(\frac{\zeta}{2} \right) \ln \left(\sin \frac{\zeta}{2} \right) \right) \kappa_{ab} \end{aligned} \quad (28)$$

where δ_a and δ_b are the angular separations between pulsars a, b and the direction of the dipole moment of the GWB.

The quadrupole correlation basis functions have the following form in the computational frame:

$$\begin{aligned} \Gamma_{2,-2} &= 0 = \Gamma_{2,-1} \\ \Gamma_{2,0} &= N \frac{1}{3} \sqrt{\frac{\pi}{5}} \left[\cos \zeta + \frac{15\beta}{4} \left(\alpha(3 + \cos \zeta) \right. \right. \\ &\quad \left. \left. + 8 \ln \left(\sin \frac{\zeta}{2} \right) \right) \right] \kappa_{ab} \\ \Gamma_{2,1} &= N \frac{1}{2} \sqrt{\frac{\pi}{15}} \sin \zeta \left[5 \cos^2 \zeta + 15 \cos \zeta - 21 \right. \\ &\quad \left. - 60 \frac{\beta}{\alpha} \ln \left(\sin \frac{\zeta}{2} \right) \right] \kappa_{ab} \\ \Gamma_{2,2} &= -N \frac{1}{4} \sqrt{\frac{5\pi}{3}} \frac{\beta}{\alpha} \left[\alpha(\cos^2 \zeta + 4 \cos \zeta - 9) \right. \\ &\quad \left. - 24\beta \ln \left(\sin \frac{\zeta}{2} \right) \right] \kappa_{ab}. \end{aligned} \quad (29)$$

The reason why the negative- m Γ_{lm} values are zero can be easily understood from the form of the negative- m real spherical-harmonic functions. These are proportional to $\sin(|m|\phi)$, which is an odd function. Since the evaluation of Γ_{lm} involves integrating over ϕ between 0 and 2π , then this will trivially yield zero.

We must rotate these computational-frame coefficients back into the cosmic frame by using the rotational transformations of the real-form spherical harmonics. Finally, for a prescribed array of pulsars with known positions, we can completely precompute the cosmic-frame correlation basis functions. These can be read into a search pipeline to be multiplied with model-dependent anisotropy coefficients, c_{lm} . The c_{lm} values will be the search parameters in such a pipeline.

IV. BAYESIAN INFERENCE

Bayes’ theorem states that the *posterior* probability density function (PDF), $p(\vec{\mu}|D, \mathcal{H})$, of the parameters $\vec{\mu}$ describing a hypothesis model \mathcal{H} , and given data D , is

TABLE I. An interpretation of the Bayes factor in determining which model is favored, as given by Jeffreys [52].

Bayes factor, \mathcal{K}	$\ln(\mathcal{K})$	Strength of evidence
<1:1	<0	Negative (supports \mathcal{H}_1)
1:1 to 3:1	0–1.1	Barely worth mentioning
3:1 to 10:1	1.1–2.3	Substantial
10:1 to 30:1	2.3–3.4	Strong
30:1 to 100:1	3.4–4.6	Very strong
>100:1	>4.6	Decisive

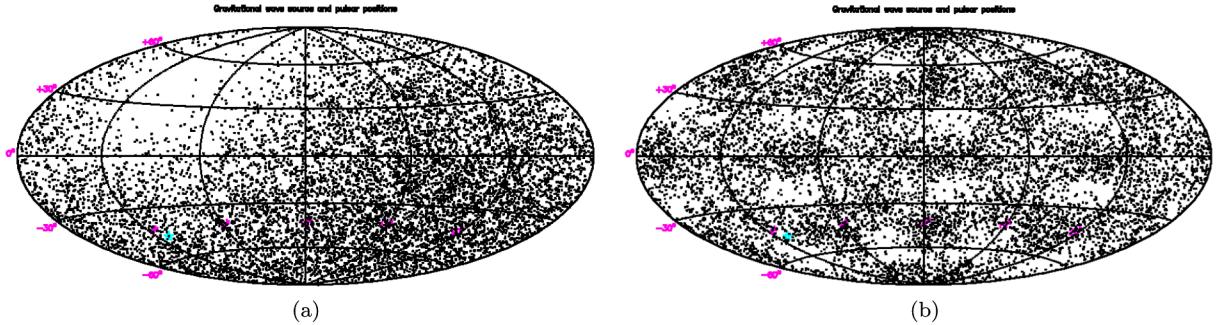


FIG. 3 (color online). Screenshots from the TEMPO2 plugins. The title of each plot is “Gravitational wave sources and pulsar positions,” with a test pulsar shown as a light-blue star and GW sources shown as black points. Lines of declination and right ascension are $[0^\circ, \pm 30^\circ, \pm 60^\circ]$ and $[4h, 8h, 12h, 16h, 20h]$, respectively. In (a) the over- and underdensity of sources in the different regions of the sky indicate the user-specified dipole direction using the `GWdipolebkrd` plugin. Likewise, the density of sources in (b) follows the user-specified multipolar structure of GW power using the `GWanisobkrd` plugin.

$$p(\vec{\mu}|D, \mathcal{H}) = \frac{p(D|\vec{\mu}, \mathcal{H})p(\vec{\mu}|\mathcal{H})}{p(D|\mathcal{H})}, \quad (30)$$

where

$$\begin{aligned} p(D|\vec{\mu}, \mathcal{H}) &\equiv \mathcal{L}(\vec{\mu}) \\ &= \text{likelihood of data given parameters,} \\ p(\vec{\mu}|\mathcal{H}) &\equiv \pi(\vec{\mu}) = \text{prior PDF of parameters,} \\ p(D|\mathcal{H}) &= \mathcal{Z} = \text{Bayesian evidence.} \end{aligned} \quad (31)$$

The Bayesian evidence, \mathcal{Z} , is the probability of the observed data given the model \mathcal{H}

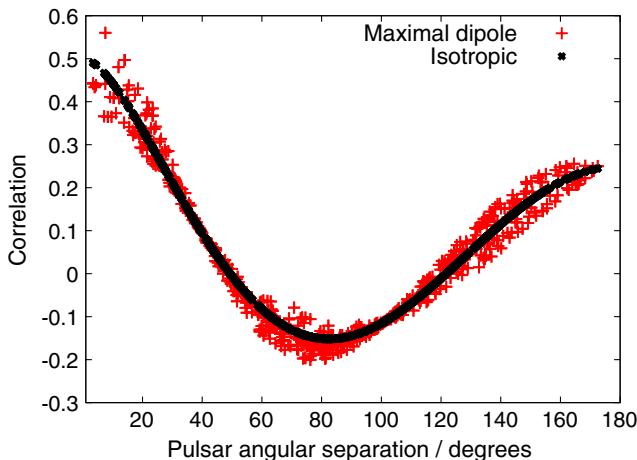


FIG. 4 (color online). Angular correlations for all pulsar pairings with nonzero angular separation in the IPTA Mock Data Challenge array of 36 pulsars. The black points indicate the angular correlation when we have an isotropic distribution of GW sources, which follows the expected Hellings and Downs curve. The red points show the angular correlations when the angular distribution of sources is proportional to $(1 + \cos \theta)$, which corresponds to a dipole in the $+z$ direction. Crucially, the scatter of these points is not noise-related, but rather encodes information about the distribution of GW sources with respect to the pulsar positions.

$$\mathcal{Z} = \int \mathcal{L}(\vec{\mu})\pi(\vec{\mu})d^N\mu. \quad (32)$$

For posterior inference within a model, \mathcal{Z} plays the role of a normalization constant and can be ignored. However, if we want to perform model selection then this evidence value becomes key. In Bayesian model comparison we compute the posterior odds ratio

$$\frac{p(\mathcal{H}_2|\vec{D})}{p(\mathcal{H}_1|\vec{D})} = \frac{p(\vec{D}|\mathcal{H}_2)p(\mathcal{H}_2)}{p(\vec{D}|\mathcal{H}_1)p(\mathcal{H}_1)} = \frac{\mathcal{Z}_2 \times p(\mathcal{H}_2)}{\mathcal{Z}_1 \times p(\mathcal{H}_1)}, \quad (33)$$

where $\mathcal{Z}_2/\mathcal{Z}_1$ is the Bayes factor, and $p(\mathcal{H}_2)/p(\mathcal{H}_1)$ is the prior probability ratio for the two competing models. The latter can often be set to 1, and we will do so in the remainder of this analysis. The posterior odds ratio is then just the Bayes factor, which is the ratio of the evidence for the two competing models. Since the evidence is the average of the likelihood over the prior volume, it automatically incorporates Occam’s razor, which states that, all else being equal, a model with fewer parameters is favored. Hypothesis \mathcal{H}_1 is chosen if the Bayes factor is sufficiently large. Jeffreys [52] gave a scale interpretation for the Bayes factor, which is shown in Table I.

V. PULSAR TIMING ANALYSIS

Repeated observation and study of a pulsar leads to a catalog of the arrival times of its pulses. After the actual radio-telescope observations are performed, the data are processed into the form of pulsar parameter files (“.par”) and timing files (“.tim”). The parameter file contains first estimates of the pulsar timing-model parameters; these parameters describe deterministic contributions to the arrival times. The vector of measured arrival times will be composed of a deterministic and a stochastic contribution (from time-correlated stochastic signals which are modeled by a random Gaussian process),

$$\vec{t}^{\text{arr}} = \vec{t}^{\text{det}} + \delta\vec{t}^{\text{gp}}. \quad (34)$$

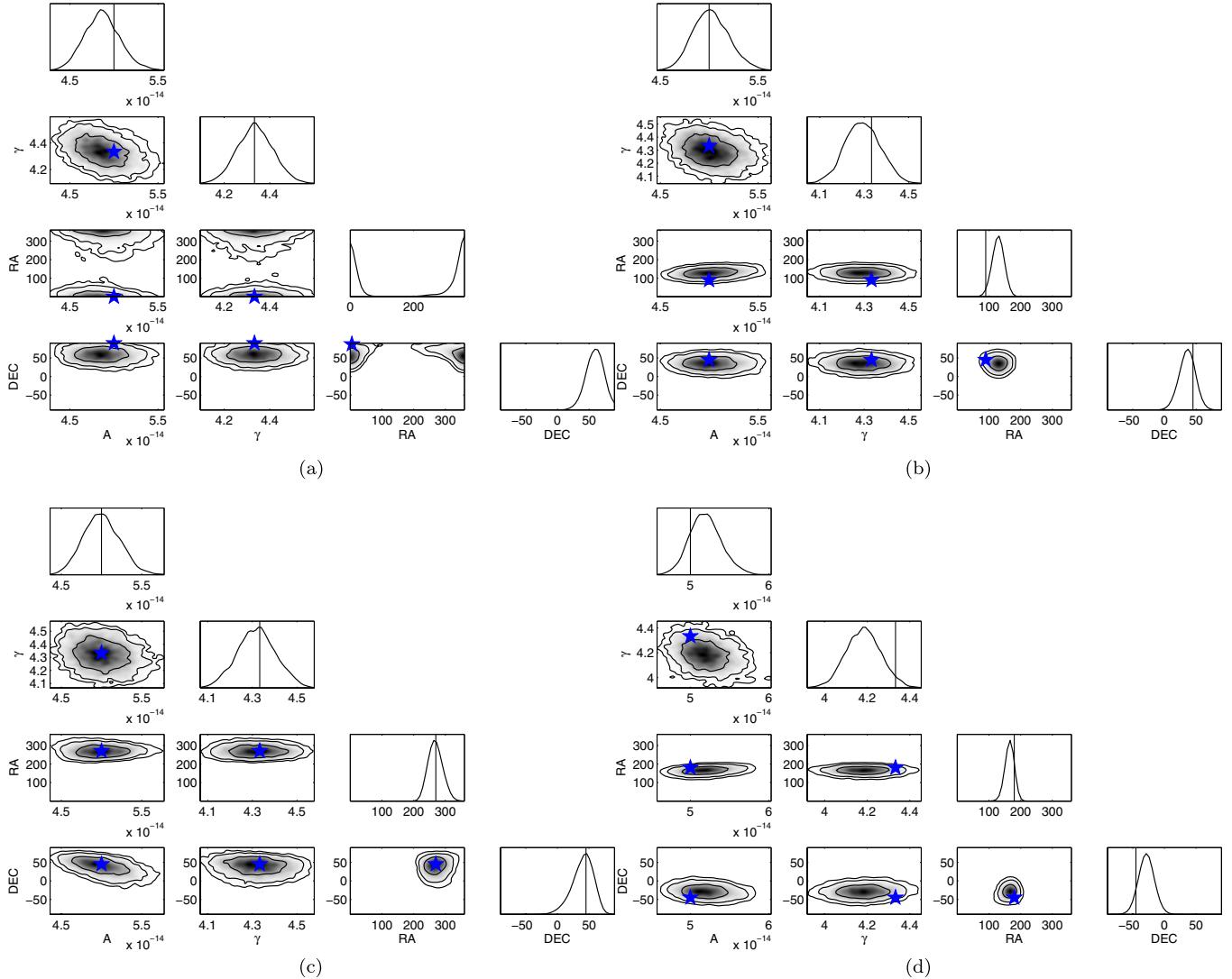


FIG. 5 (color online). Recovered posterior probability distributions for a Bayesian parameter estimation of four different dipolar GWBs in simulated PTA data. The anisotropy formalism is able to recover the direction of the injected dipole, as well as the properties of the underlying GWB spectrum. The strain spectrum is a power law in all cases, with $(A = 5 \times 10^{-14}, \gamma = 13/3)$, and the injected dipole direction is (a) $(RA = 0^\circ, DEC = 90^\circ)$, (b) $(RA = 90^\circ, DEC = 45^\circ)$, (c) $(RA = 270^\circ, DEC = 45^\circ)$, (d) $(RA = 180^\circ, DEC = -45^\circ)$. There is negligible correlation between the spectral index γ and the dipole direction, which indicates that we can fix γ to its injected value without biasing the anisotropy constraints. Fixing γ affords the analysis large accelerations. We assume a maximal dipole in the analysis of these data sets; i.e., the relative amplitude of the isotropic and dipole components is fixed to 1 and not included as a model parameter.

The stochastic process is characterized by its autocorrelation,

$$C_{ij} = \langle \delta t_i^{\text{rgp}} \delta t_j^{\text{rgp}} \rangle, \quad (35)$$

where the elements of the covariance matrix are parametrized by a set of parameters, $\vec{\phi}$. Using the Wiener-Khinchin theorem, we can then define the autocorrelation as the Fourier transform of the power spectral density,

$$C(\tau_{ij}) = \int_0^\infty S(f) \cos(f\tau_{ij}) df, \quad (36)$$

where $\tau_{ij} = 2\pi|t_i - t_j|$, and $S(f)$ is the power spectral density of the time series δt^{rgp} . An analytic expression

for the autocorrelation of a time series influenced by an underlying power-law power spectral density is given in van Haasteren *et al.* [35], and is used in the following.

We describe the stages of pulsar timing analysis as either *offline* or *online*, depending on whether they are performed before or during the sampling of parameter space via likelihood evaluations.

A. Processing raw arrival times (offline)

The .par and .tim files are fed to the TEMPO2 software package [43–45] which processes the raw arrival times. Vectors of “prefit” timing residuals are computed using

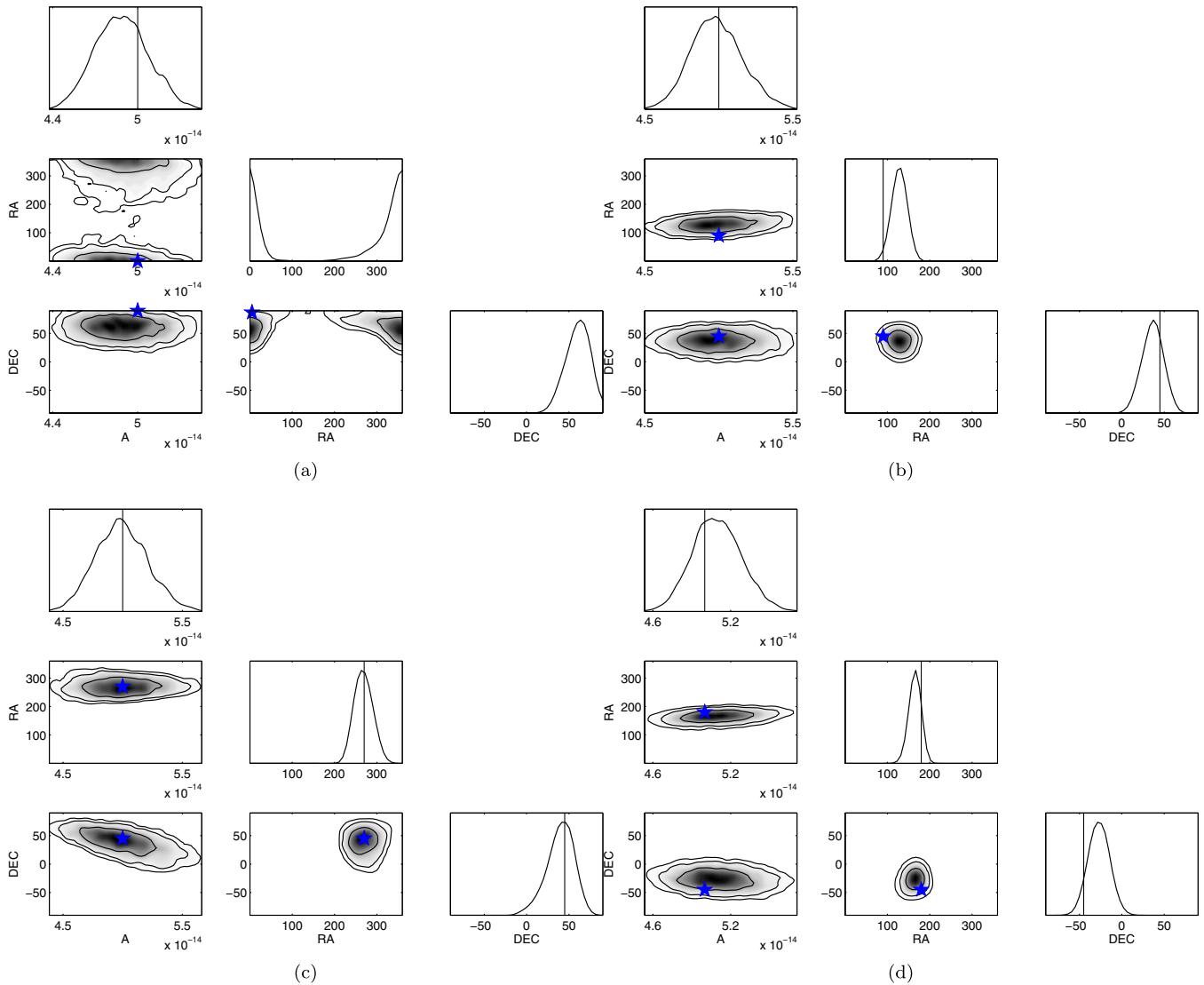


FIG. 6 (color online). We repeat the analysis of the dipole-direction recovery, as in Fig. 5, but this time fix γ to $13/3$. As expected, we are able to recover A and the (RA, DEC) of the injected dipole without bias.

the first guesses, $\beta_{0,i}$, of the “ m ” timing-model parameters from the .par files. This first guess is usually precise enough to allow a linear approximation to be used in the TOA fitting procedure. The postfit timing residuals are then given by

$$\delta\vec{t} = \delta\vec{t}^{\text{pref}} + M\vec{\xi}, \quad (37)$$

where $\delta\vec{t}^{\text{pref}}$ are the prefit timing residuals (length n), $\vec{\xi}$ is the vector of deviations from the prefit parameters (length m) defined as $\xi_a = \beta_a - \beta_{0,a}$, and M is the $(n \times m)$ “design matrix,” describing how the residuals depend on the timing-model parameters. TEMPO2 does not take into account the possible time-correlated stochastic signal in the TOAs, but performs a weighted least-squares fit for the timing-model parameter values. Hence it is possible that

some of the time-correlated stochastic signal is removed by this fitting procedure, which is undesirable.

The TEMPO2 analysis provides output residuals and the design matrix, M . The design matrix describes the dependence of the timing residuals on the timing-model parameters. The output residuals form the input data vector for further study.

B. The time-domain likelihood (online)

We now use the TEMPO2 output residuals to search for the presence of any correlated stochastic signal affecting the pulse arrival times. We assume that the part of the stochastic signal removed by the fitting procedure is small, so that the TEMPO2 output residuals are related linearly to the stochastic contribution to the residuals [53]

$$\delta\vec{t} = \delta\vec{t}^{\text{rgp}} + M\vec{\xi}, \quad (38)$$

TABLE II. Evidence differences between an anisotropic and isotropic model when analyzing data sets which have dipole anisotropies injected in various different directions. In each case we favor the model which accounts for this anisotropy rather than the isotropic model. This is true even when the direction of the dipole is searched over, which adds an extra two dimensions to the parameter space. The Bayes factor varies little when the spectral index $\gamma \equiv 3 - 2\alpha$ is fixed to 13/3, which is the expected value for a background composed of inspiraling SMBHBs.

Dipole direction/degrees	$\Delta \ln(Z) = \ln(Z_{\text{true}}) - \ln(Z_{\text{iso}})$	
	$\gamma = 13/3$	γ varied
RA = 0.0 DEC = 90.0	7.0	7.2
RA = 90.0 DEC = 45.0	10.5	10.3
RA = 270.0 DEC = 45.0	4.4	4.4
RA = 180.0 DEC = -45.0	7.3	7.5

where, in this case, $\delta\vec{t}$ refers to the output residuals from TEMPO2. We note that the $\vec{\xi}$ appearing in this equation is different from that appearing in Eq. (37).

The stochastic timing residuals, $\delta\vec{t}^{\text{reg}}$, arise from a time-correlated stochastic process with covariance matrix C [see Eq. (35)]. This covariance matrix may contain contributions from the GWB, white noise from TOA errors, and possibly red-timing noise which is uncorrelated between different pulsars. The likelihood of measuring postfit residuals $\delta\vec{t}$, given stochastic parameters $\vec{\phi}$, and marginalizing with flat priors over all timing-model parameters is [53]

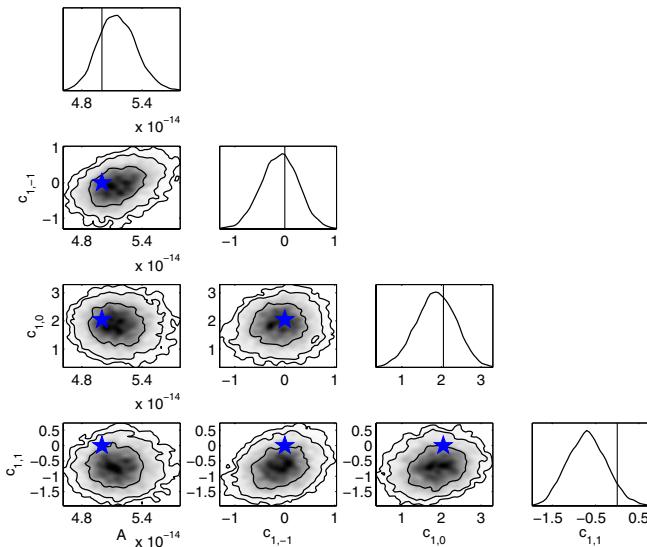


FIG. 7 (color online). Results of a search over all dipole anisotropy coefficients, with the assumption relaxed that the relative amplitude of the dipole and isotropic components is maximal. The injected anisotropy was maximal dipole with $c_{0,0} = 2\sqrt{\pi}$ and $c_{1,0} = 2\sqrt{\pi}/3$, corresponding to a dipole direction of RA = 0.0°, DEC = 90°. The anisotropic-search pipeline fully recovered all details of the injected anisotropy, as well as the amplitude of the GWB characteristic strain spectrum, A .

$$\begin{aligned} \mathcal{L}(\delta\vec{t}|\vec{\phi}) &= \frac{1}{\sqrt{(2\pi)^{n-m} \det(G^T G)}} \\ &\times \exp\left(-\frac{1}{2}\delta\vec{t}^T G(G^T G)^{-1} G^T \delta\vec{t}\right), \end{aligned} \quad (39)$$

where G is the matrix constructed from the final $(n - m)$ columns of the matrix U in the singular value decomposition of the design matrix, $M = U\Sigma V^*$. The matrix G can be precomputed and stored in memory for use in each likelihood calculation (hence the computation of G is offline). Equation (39), with appropriate priors on the stochastic parameters, not only provides a robust, unbiased Bayesian framework for the search for correlated signals in PTAs, but also incorporates a lossless data compression through the G -matrix projection [54].

In the case of multiple pulsars, where we have processed them independently with TEMPO2, the total covariance matrix, total G matrix and total residual vector are given by

$$\begin{aligned} C &= \begin{pmatrix} C_{11} & C_{12} & \dots \\ C_{21} & C_{22} & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}, \\ G &= \begin{pmatrix} G_1 & & \\ & G_2 & \\ & & \ddots \end{pmatrix}, \\ \delta\vec{t} &= \begin{pmatrix} \delta\vec{t}_1 \\ \delta\vec{t}_2 \\ \vdots \end{pmatrix}, \end{aligned}$$

where C_{ab} is the autocovariance matrix between pulsars a and b , G_a are the individual pulsar timing-model marginalization matrices and $\delta\vec{t}_a$ are the individual pulsar residual vectors. We can split C_{ab} into contributions from various stochastic sources. In this paper we consider only the stochastic influence due to a GWB (which has correlation Γ_{ab} between pulsars a and b), TOA error bars (white and uncorrelated between different pulsars), and intrinsic pulsar red noise (uncorrelated between different pulsars). So, the covariance between the i th residual of pulsar a and the j th residual of pulsar b is

$$C_{(ai)(bj)} = C_{(ai)(bj)}^{\text{GW}} + C_{(ai)(bj)}^{\text{TOA}} + C_{(ai)(bj)}^{\text{RN}}, \quad (40)$$

where C_{ab}^{TOA} , C_{ab}^{RN} and C_{ab}^{GW} are diagonal, block-diagonal and block-symmetric respectively.

The spectral density of induced timing residuals for the intrinsic red noise is taken to have the same form as a GWB, with amplitude A_{red} and slope γ_{red} .

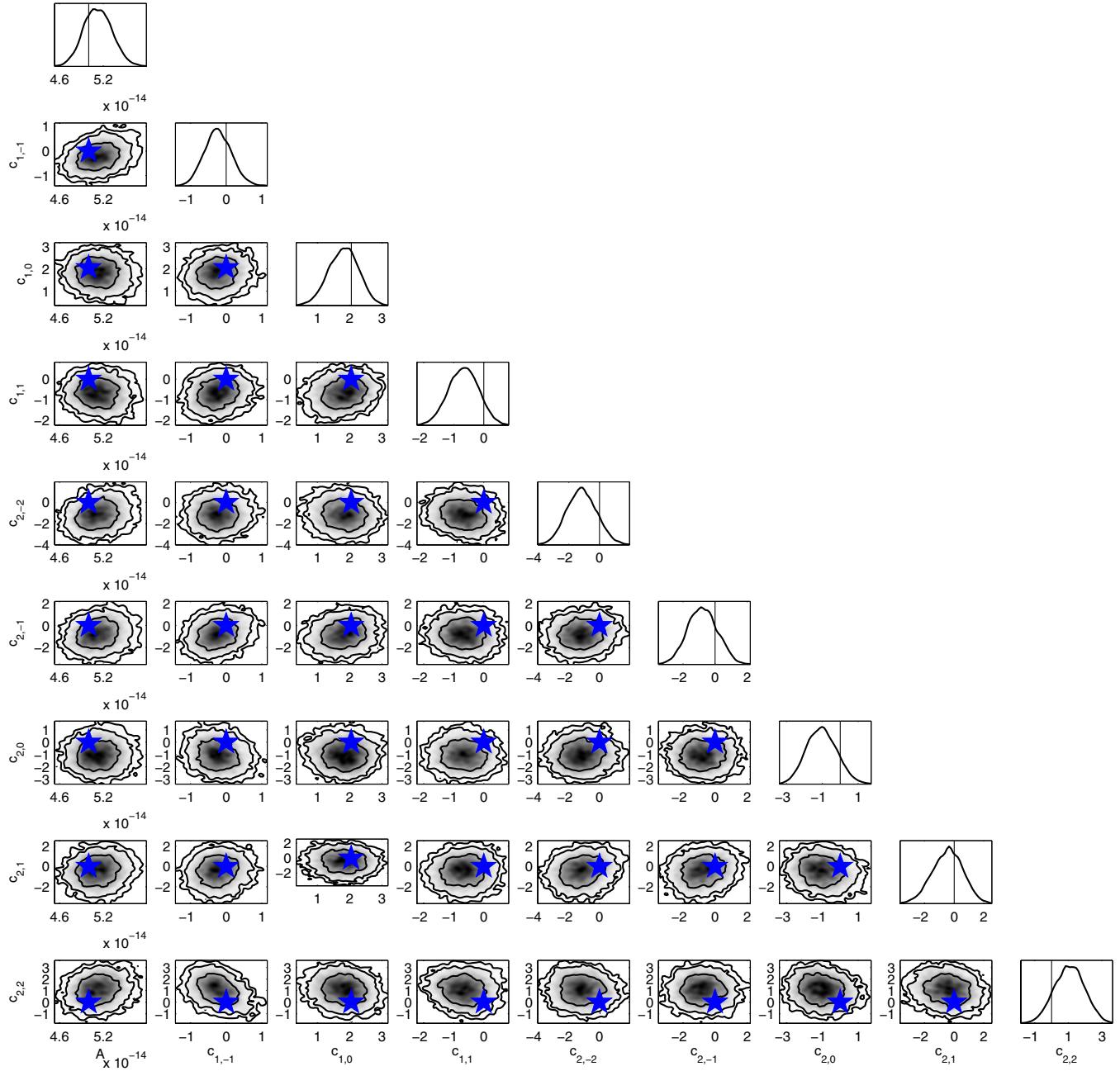


FIG. 8 (color online). A search over all anisotropy coefficients, c_{lm} , up to, and including, quadrupole order. The injected anisotropy was maximal dipole with $c_{0,0} = 2\sqrt{\pi}$ and $c_{1,0} = 2\sqrt{\pi}/3$, corresponding to a dipole direction of RA = 0.0°, DEC = 90°. As can be seen, the anisotropic-search pipeline fully recovered all details of the injected anisotropy, as well as the amplitude of the GWB characteristic strain spectrum, A .

VI. SIMULATING AN ANISOTROPIC BACKGROUND

To simulate anisotropy in a GWB we have developed a suite of new TEMPO2 plugins. These permit a user to define not only the amplitude and spectral index of the underlying characteristic strain spectrum of the background, as in the GWbkgd plugin, but also the angular distribution of the sources comprising the background. We can also define the polarization of the sources comprising the background,

which we have not limited to the Einsteinian polarization states.

To generate an anisotropic background we define a probability density function for the placement of sources on the sky,

$$\begin{aligned} \mathcal{P}(\theta \in [\Theta, \Theta + d\theta], \phi \in [\Phi, \Phi + d\phi]) \\ = P(\Theta, \Phi) d\theta d\phi \end{aligned} \quad (41)$$

where we enforce $P(\Theta, \Phi) \geq 0$.

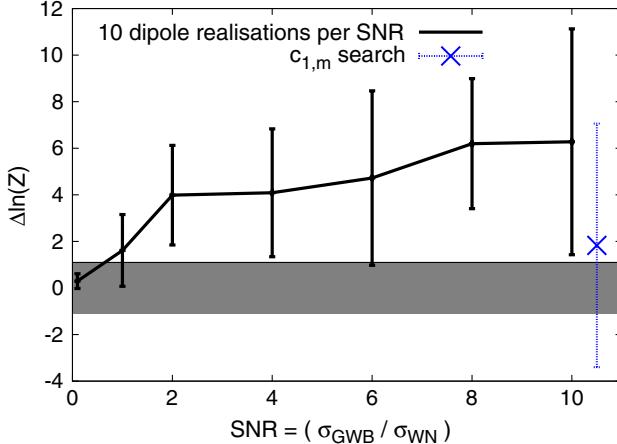


FIG. 9 (color online). Evidence for anisotropy versus isotropy as a function of the amplitude of the GWB. We generate many data sets with dipole anisotropy such that the angular distribution of GW sources is $\propto (1 + \cos \theta)$. We vary the SNR of the GWB by fixing the white-noise level and changing the amplitude of the characteristic strain spectrum A , where for each SNR value we generate 10 data sets. The SNR range 0.1–10 corresponds to varying A between 5×10^{-16} – 5×10^{-14} . We fix γ to its fiducial value of $13/3$, and evaluate the Bayesian evidence for models which assume the correct injected anisotropy (c_{lm} fixed) and isotropy. The shaded region indicates the range of Bayes factors for which we cannot tell the difference between anisotropic and isotropic models. Since the injected noise is white-uncorrelated, and we compress to preserve low-frequency information, even at $\text{SNR} = 0.1$ the GWB is loud enough to be detected with substantial evidence (see text for details). The blue, dotted point is a calibration point for more general searches over $\{A, c_{1,m}\}$ at $\text{SNR} = 10$, but offset on the x axis for ease of viewing (see text for details).

The background structure is user-defined, and is set by entering a multipolar decomposition of the sky-location PDF, as well as the usual dimensionless background amplitude, A , and strain-spectrum slope, α . We now give a brief description of the individual plugins, which are freely available.

- (i) **GWdipolebkrd:** Generates a background with a dipolar angular distribution of the GW power. The user specifies the multipole coefficients, $\{c_{00}, c_{1-1}, c_{10}, c_{11}\}$, or the direction of the dipole moment.
- (ii) **GWanisobkrd:** Generates a background with arbitrary angular distribution of the GW power. The user specifies the multipole coefficients, $\{c_{lm}\}$.
- (iii) **GWgeneralbkrd:** Generates an isotropic background composed of source populations with different GW polarizations, i.e. TT (GR modes), ST (scalar transverse, or *breathing*, mode), SL (scalar longitudinal) or VL (vector longitudinal). User specifies A and α for each polarization.
- (iv) **GWgeneralanisobkrd:** Same as GWgeneralbkrd, but with an arbitrary angular distribution of sources

for each polarization, specified by a set of multipole coefficients, $\{c_{lm}^A\}$, for each polarization state A .

Figure 3 shows screenshots displaying the distribution of sources comprising several examples of user-specified backgrounds. These plugins and explanations of their usage can be found at <https://github.com/jonathangair/tempo2-anisotropy-plugins>.

VII. RESULTS

We generated data sets similar to those used for Open Dataset 1 of the first IPTA Mock Data Challenge [39]. They contained 36 pulsars distributed across the sky and timed fortnightly for 5 years. The injected signal was a GWB consistent with a background of inspiraling SMBHBs (i.e. $\gamma = 13/3$), and the pulsars were timed to 100 ns accuracy.

For a full data set this amounts to 4680 timing residuals. Even with highly tuned libraries and multithreading, the required matrix operations are very costly. We employ a high-fidelity data-compression technique developed by van Haasteren [54] for the purposes of throwing away unwanted high-frequency information and retaining as much information about the GWB as possible, which has a steep, red spectrum.

The data compression is based on diagonalizing the expected signal covariance matrix, and determining which basis vectors have the greatest contribution to the Fisher information. A set of reduced basis vectors are then found which capture as much information about the signal of interest as possible. In the lexicon of [54], we demand a minimum fidelity of 0.99. With $A = 5 \times 10^{-14}$, $\sigma_{WN} = 100$ ns and 0.99 fidelity we can compress the number of residuals by approximately a factor of 5. Matrix multiplications and inversions are $\mathcal{O}(n^3)$ operations, such that we achieve significant acceleration.

All evidence evaluation and parameter estimation was performed using nested sampling [55] with the MULTINEST package [56,57].

A. Dipole injections

The nature of a dipole anisotropy in a GW-source distribution can be seen in Fig. 3(a), where we see an overdensity of sources in a particular sky location, as well as a paucity of sources in the opposite sky direction. Figure 4 shows the angular correlations of pulsar pairings with non-zero angular separation, Γ_{ab} , for all pulsars in our array. We also show the isotropic angular correlation, which is the familiar Hellings and Downs curve [40] and only depends on the angular separation of the pulsars. However if we have a dipole anisotropy in the angular distribution of GW sources on the sky, then the resultant pulsar angular correlations depend not only on the angular separation between the pulsars, but also on the absolute positions of the pulsars relative to the GW-source anisotropy. Hence pulsar pairs which may have equal angular separations will not necessarily have equal angular correlations if there is anisotropy

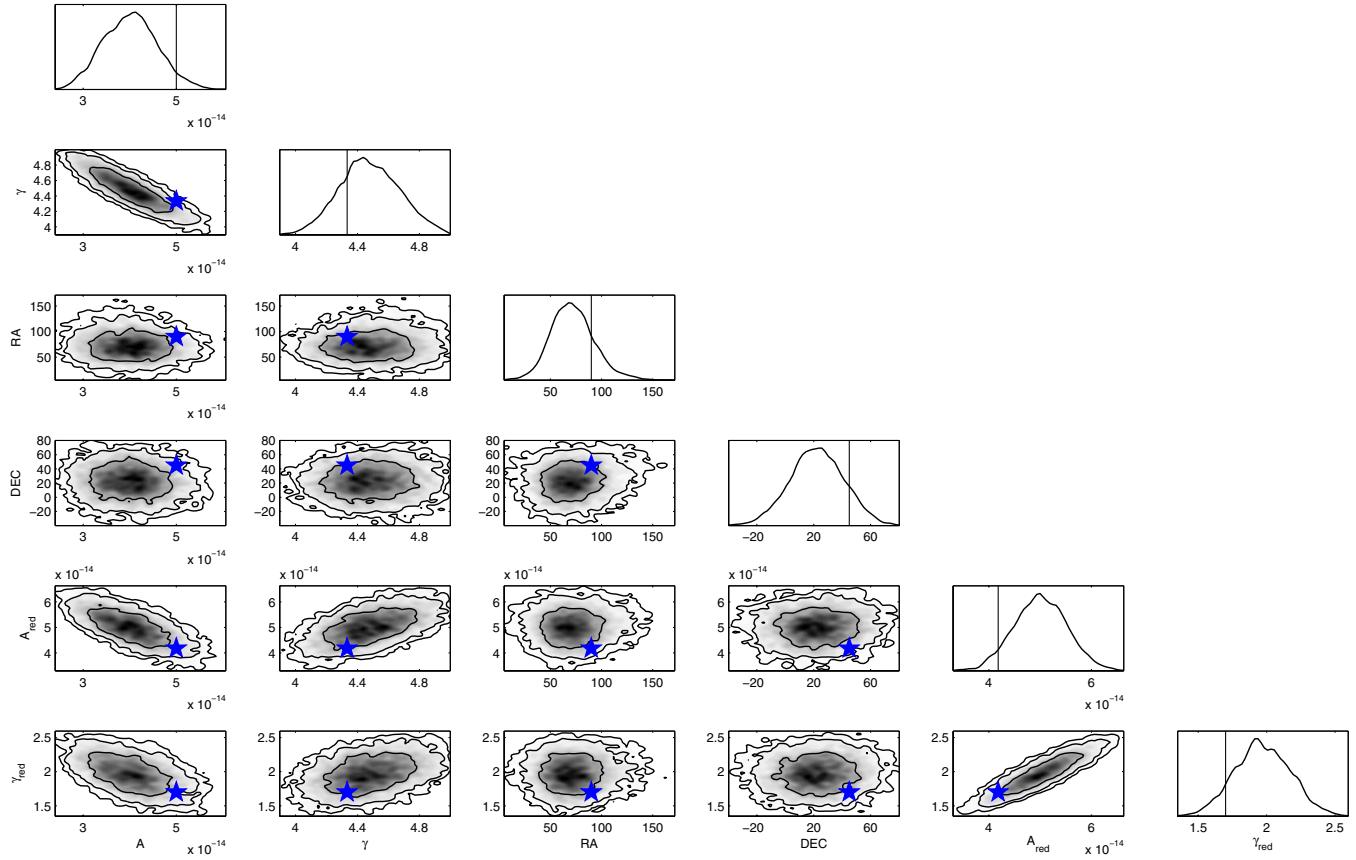


FIG. 10 (color online). An example of anisotropy reconstruction when uncorrelated red noise is present in the data set. The anisotropy is maximal dipole, with direction ($RA = 90^\circ$, $DEC = 45^\circ$). The injected GWB parameters are ($A = 5 \times 10^{-14}$, $\gamma = 13/3$). The red noise is uncorrelated between different pulsars, but shares a common spectrum, with ($A_{\text{red}} = 4.2 \times 10^{-14}$, $\gamma_{\text{red}} = 1.7$), such that we have an effective SNR of ~ 4.2 . In this data set, the $\ln(\mathcal{B})$ for anisotropic versus isotropic was ~ 4.4 , which is borderline decisive evidence for anisotropy.

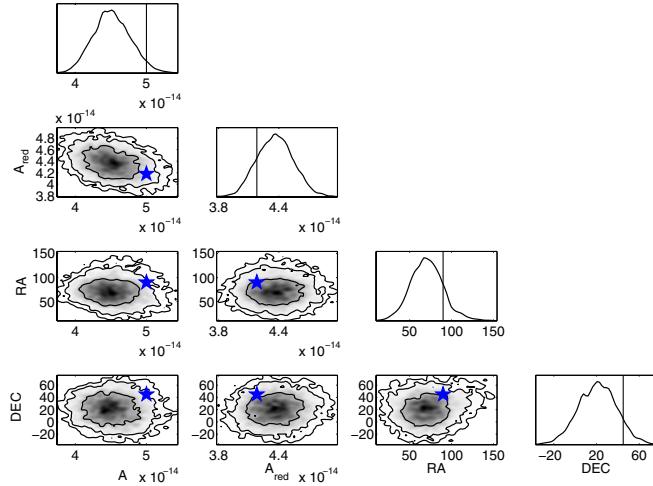


FIG. 11 (color online). Reconstructing the amplitudes of the GWB and the uncorrelated red-noise spectrum when their spectral slopes are fixed at the injected values of $13/3$ and 1.7 respectively. The data set is the same as in Fig. 10. By fixing the spectral-slope values, we can accelerate the analysis without biasing the dipole-direction reconstruction. Furthermore, the value of $\ln(\mathcal{B})$ in this 4D case is only ~ 0.8 larger than in the full 6D scenario.

in the GW-source population. Figure 4 shows the angular correlation when the angular distribution of sources has the maximum dipole amplitude that satisfies Eq. (41) and the dipole direction is along the $+z$ axis. Such a maximal dipole distribution is proportional to $(1 + \cos \theta)$.

We see that a dipole anisotropy induces a small deviation away from the isotropic Hellings and Downs curve, which encodes information about the angular structure of the GWB.

We generated data sets with dipole anisotropies in various different directions on the sky, and used the formalism of Sec. III B to constrain the properties of the underlying GWB spectrum, as well as the direction of the dipole. For the latter, the alternative phrasing of dipole anisotropy of Anholm *et al.* [46] is useful, since the dipole correlation is simply expressed as a function of the pulsar-pulsar separations and the pulsar-dipole separations.

The results of these injections are shown in Figs. 5 and 6, where we show the recovered posterior probability distributions for the search parameters. These posterior probability distributions show the ability of a Bayesian analysis to constrain the direction of a dipole in the GW-source population. Figure 5 shows a search in which the spectral

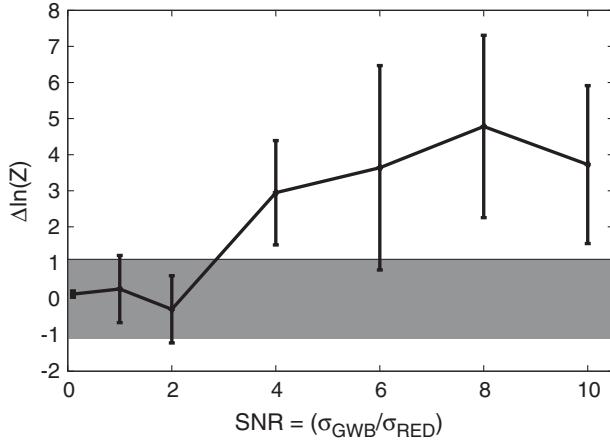


FIG. 12. As in Fig. 9, but now for data sets containing red noise. We generate many data sets with dipole anisotropy such that the angular distribution of GW sources is $\propto 1 + \cos \theta$. We vary the SNR of the GWB by fixing the characteristic strain spectrum amplitude, A , at 5×10^{-14} and changing the amplitude of the red-noise level, where for each SNR value we generate 5 data sets. The SNR range 0.1–10 corresponds to varying A_{red} between $\sim 1.8 \times 10^{-12}$ – 1.8×10^{-14} . We fix the slopes of the power spectral density of the induced residuals due to a GWB and uncorrelated red noise to the injected values of 13/3 and 1.7, respectively, and evaluate the Bayesian evidence for models which assume a maximal-dipole anisotropy (direction searched over) and isotropy. The shaded region indicates the range of Bayes factors for which we cannot tell the difference between anisotropic and isotropic models.

index, γ , was included as a parameter, whereas Fig. 6 fixes γ to the injected value of 13/3. Fixing γ accelerates the analysis since we can precompute various matrices and avoid some expensive matrix multiplications. As we can see from Fig. 5 there is negligible correlation between γ and the dipole direction, so fixing γ does not bias the reconstruction.

The only inconsistency in the reconstructed direction was found when the dipole direction was ($\text{RA} = 0.0^\circ$, $\text{DEC} = 90.0^\circ$). This arose because we sample uniformly in $\cos \theta$ so there is little prior weight at high declinations. There is also the question of the significance of accounting for the anisotropy. In other words, does the anisotropic model provide a significantly better fit to the data than a standard isotropic search would? We answer this question by inspecting the difference in $\ln(Z)$ between models which account for anisotropy, and for models which assume isotropy. The results are shown in Table II, where we see that in all cases the true model is favored, despite it having two extra dimensions compared to the isotropic model. However, we emphasize that the large Bayes factors are partially the result of our injections being in the very strong-signal regime.

To test the robustness of the *generalized* anisotropy formalism, we repeated the analysis of a maximal-dipole data set in the direction ($\text{RA} = 0^\circ$, $\text{DEC} = 90^\circ$), but attempted a reconstruction of all the $c_{1,m}$ coefficients. We sample the

anisotropy coefficients uniformly over the range $\in [-5, 5]$. Since this is a strong anisotropic signal, and we expect to be likelihood-dominated, we adopt this uniform sampling range as a prior over the anisotropy coefficients here, and defer a proper discussion of a physically motivated prior required for the analysis of real data until Sec. VII C.

We fix $c_{0,0}$ to $2\sqrt{\pi}$, and absorb the variation of the isotropic power into the overall amplitude A . So, in this sense, we are searching for deviations from isotropy, rather than highly anisotropic backgrounds. Furthermore, by freely searching over $c_{1,m}$ we are removing our assumption of the anisotropy being maximal dipole. The results are shown in Fig. 7, where we see that the recovered dipole coefficients are completely consistent with the injected values of $\{c_{1,-1} = 0, c_{1,0} = 2\sqrt{\pi}/3, c_{1,1} = 0\}$, which correspond to a maximal dipole in the $+z$ direction. The results of an analysis where we relax the assumption of *dipole* anisotropy, and perform a search over all anisotropy coefficients up to and including quadrupole, are shown in Fig. 8. The results, as before, are consistent with the injected anisotropy; however the increased dimensionality widens the posteriors such that we could conceivably be consistent with a large range of anisotropic distributions.

We now investigate the degree to which an isotropic search is suboptimal by generating many dipole-anisotropy data sets with different GWB amplitudes. The white-noise level is fixed at 100 ns, such that we are effectively varying the signal-to-noise ratio (SNR) in the data sets. We parametrize SNR in the following way:

$$\text{SNR} \sim \frac{\sigma_{\text{GW}}}{\sigma_{\text{WN}}} \quad (42)$$

where

$$\begin{aligned} \sigma_{\text{GW}} &\sim 1.37 \times \left(\frac{A}{1 \times 10^{-15}} \right) \left(\frac{T}{\text{yr}} \right)^{5/3} \text{ ns}, \\ \text{if } \gamma &= 13/3 \\ \sigma_{\text{WN}} &= 100 \text{ ns} \end{aligned} \quad (43)$$

and σ_{GW} is determined from an evaluation of the postfit covariance function for power-law spectral densities. An estimate of this is obtained analytically by projecting the prefit covariance function into a new basis, which replicates the effect of fitting for quadratics. Given that the quadratic-spindown fitting has the dominant effect on the covariance matrix, Eq. (43) provides a good estimate of σ_{GW} for given A (and vice versa) [53,54,58].

Varying the amplitude of the GWB between 5×10^{-16} – 5×10^{-14} gives an effective SNR span of 0.1–10. At each SNR in this range we generated 10 data set realizations and evaluated the difference in $\ln(Z)$ values for an anisotropic and an isotropic model. This was done with the GWB correlations fixed to those corresponding to the injected level of anisotropy, which in this case was the maximal dipole in the $+z$ direction described earlier.

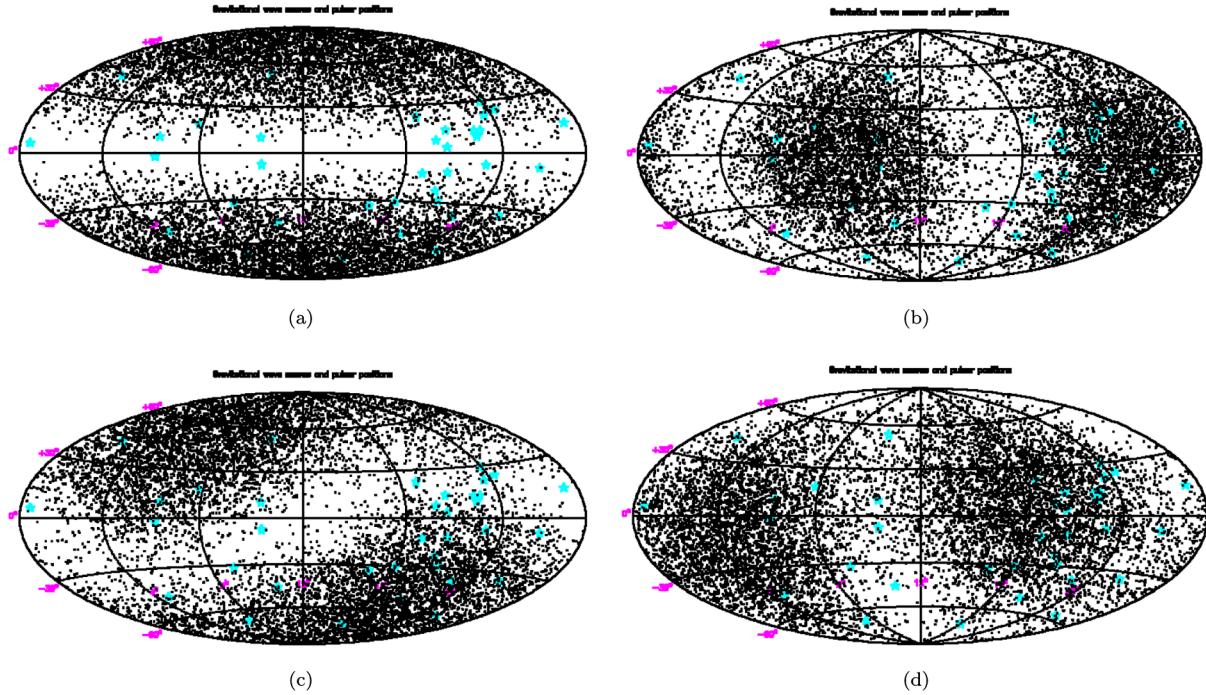


FIG. 13 (color online). Skymap screenshots from the GWanisobkgrd TEMPO2 plugin showing the quadrupolar angular distribution of GW sources constituting various GWB realizations. The title of each plot is “Gravitational wave sources and pulsar positions,” with pulsars shown as light-blue stars and GW sources shown as black points. Lines of declination and right ascension are $[0^\circ, \pm 30^\circ, \pm 60^\circ]$ and $[4h, 8h, 12h, 16h, 20h]$, respectively. The anisotropy coefficients for each data set are (a) $\{c_{0,0} = 2\sqrt{\pi}; c_{2,0} = 4\sqrt{\pi/5}\}$, (b) $\{c_{0,0} = 2\sqrt{\pi}; c_{2,-2} = -0.9c_{0,0}; c_{2,0} = -0.3c_{0,0}\}$, (c) $\{c_{0,0} = 2\sqrt{\pi}; c_{2,-1} = -0.8c_{0,0}; c_{2,0} = 0.5c_{0,0}; c_{2,2} = -0.3c_{0,0}\}$, (d) $\{c_{0,0} = 2\sqrt{\pi}; c_{2,-2} = 0.56c_{0,0}; c_{2,-1} = 0.20c_{0,0}; c_{2,0} = -0.32c_{0,0}; c_{2,1} = 0.27c_{0,0}; c_{2,2} = -0.40c_{0,0}\}$.

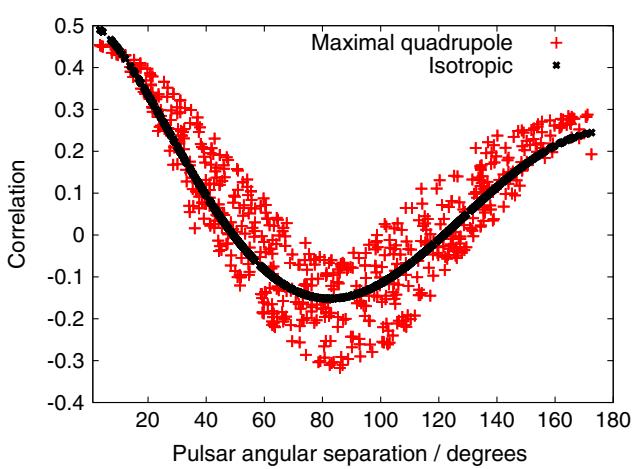


FIG. 14 (color online). Angular correlations for all pulsar pairings with nonzero angular separation in the IPTA Mock Data Challenge array of 36 pulsars. The black points indicate the angular correlation when we have an isotropic distribution of GW sources, which follows the expected Hellings and Downs curve. The red points show the angular correlations when the angular distribution of sources follows $\propto 1 + (3\cos^2\theta - 1)$, which corresponds to a quadrupole anisotropy. The quadrupole clearly induces a larger deviation from the Hellings and Downs curve than the dipole anisotropy in Fig. 4.

These results are shown in Fig. 9. We note that this figure does not show the true odds ratio for the anisotropic versus isotropic model, since the degrees of freedom corresponding to the dipole direction have not been included in the parameter space for the anisotropic model. Instead, this figure indicates a best-case scenario for when the evidence would start to favor the anisotropic model. Also, one should note that because the injected noise is white-uncorrelated, and we have compressed to preserve information about a steep red-noise process, even at $\text{SNR} = 0.1$ the GWB is loud enough to be detected with substantial evidence. At this SNR the largest noise-versus-(isotropic-)GWB log-Bayes factor is 0.03, with a mean and median of -6.30 and -2.62 , respectively. So (in this parametrization of SNR), at $\text{SNR} = 0.1$ we can detect a GWB with angular correlation of *approximate* Hellings and Downs form, but the background is not yet loud enough for anisotropy-induced deviations to be visible.

We see that in these idealized circumstances the anisotropic model is marginally favored for $\text{SNRs} > \sim 1$ and decisively favored for $\text{SNRs} > \sim 5$. Including the dipole direction parameters (or searching over $c_{1,m}$) in the parameter space reduces the anisotropic evidence by several and so in practice $\text{SNRs} > \sim 5\text{--}10$ would be required before the presence of anisotropy would become

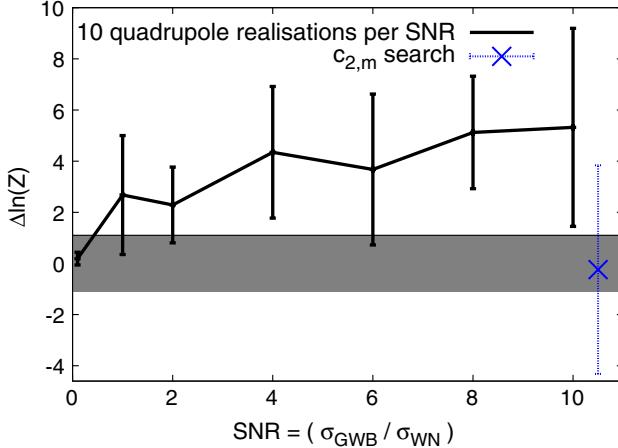


FIG. 15 (color online). As in Fig. 9, but now for data sets containing a maximal quadrupole anisotropy such that the angular distribution of GW sources is $\propto 1 + (3\cos^2\theta - 1)$. We vary the SNR of the GWB by fixing the white-noise level and changing the amplitude of the characteristic strain spectrum A , and for each SNR value we generate 10 data sets. The SNR range 0.1–10 corresponds to varying A between 5×10^{-16} – 5×10^{-14} . We fix γ to its fiducial value of 13/3, and evaluate the Bayesian evidence difference between a model that assumes the correct injected anisotropy (c_{lm} fixed) and an isotropic model. The shaded region indicates the range of Bayes factors for which we cannot tell the difference between anisotropic and isotropic models. Since the injected noise is white-uncorrelated, and we compress to preserve low-frequency information, even at $\text{SNR} = 0.1$ the GWB is loud enough to be detected with substantial evidence (see text for details). The blue, dotted point is a calibration point for more general searches over $\{A, c_{lm}\}$ at $\text{SNR} = 10$, but offset on the x axis for ease of viewing (see text for details).

clearly visible to an analysis of this type (see also Fig. 12 below). We provide a calibration point at $\text{SNR} = 10$, where we search over $\{A, c_{1,m}\}$ for the 10 data set realizations at this SNR. We see that the increased dimensionality of the search penalizes the recovered Bayes factors; however the average log-Bayes factor implies that the presence of anisotropy may become clearly visible at SNRs greater than 10.

The data sets described so far have contained only white noise. Now we generate data sets with an additional red-noise component uncorrelated between different pulsars. The red noise shares a common spectrum, but with a different realization for each pulsar. The spectral index was taken to match that of the IPTA Open3 data set, which was $\gamma_{\text{red}} = 1.7$. As before, we include an uncorrelated white-noise component of 100 ns in each pulsar. Following Eq. (43), we parametrize the SNR as

$$\text{SNR} \sim \frac{\sigma_{\text{GW}}}{\sigma_{\text{RED}}} \sim 0.422 \times \left(\frac{A}{A_{\text{red}}} \right) \times \left(\frac{T}{\text{yr}} \right)^{79/60} \quad (44)$$

where the form of σ_{RED} follows from the more general form of the rms residual induced by a stochastic process with a power-law spectrum of arbitrary slope [53].

The analysis is more expensive than a GWB-only analysis, since we must also characterize and constrain the red-noise properties. Furthermore, the red noise has a shallower spectrum, such that our residuals contain more high-frequency noise, and we cannot compress to the same degree as in the case of a GWB-only data set. Figure 10 shows the results of a $\text{SNR} \sim 4$ analysis. We see that at this SNR we can characterize the GWB, red-noise and anisotropy properties. In fact, for this data set realization, the value of $\ln(\mathcal{B})$ for anisotropy versus isotropy was ~ 4.4 , which is borderline decisive evidence for anisotropy. Since the analysis is quite computationally expensive, and we wish to perform a more systematic study over many background realizations, we repeat this analysis but fix γ and γ_{red} to the injected values of 13/3 and 1.7, respectively. Fixing these spectral indices accelerates the analysis, does not bias the parameter reconstruction (see Fig. 11), and only raises the value of $\ln(\mathcal{B}) = \Delta \ln(Z)$ by ~ 0.8 .

We perform a similar study to Fig. 9, generating 5 data sets per SNR, each with common GWB properties ($A = 5 \times 10^{-14}$, $\gamma = 13/3$, $\gamma_{\text{red}} = 1.7$, but varying the red-noise amplitude, A_{red} to simulate the different SNRs. The injected anisotropy is maximal dipole in the ($\text{RA} = 90^\circ$, $\text{DEC} = 45^\circ$) direction in all cases. We search over the parameters $\{A, A_{\text{red}}, \text{RA}, \text{DEC}\}$. The results are shown in Fig. 12, where we see that the presence of uncorrelated red noise in the data sets can be accounted for, but its presence means that the SNR required to identify the presence of anisotropy in the data appears somewhat higher. However, in this case we include a search over the dipole direction in the evaluation of the evidence, so these results are much closer to what might be achievable in practice (although fixing the slope of the red-noise spectrum in particular might not be a valid assumption). Figure 12 suggests that even in the presence of pulsar red noise the evidence for anisotropy will become apparent for SNRs around 3 and will become decisive for SNRs of ~ 6 .

For the remainder of this analysis we ignore red noise and inject only uncorrelated white noise in the form of TOA error bars, but expect that the effect of red noise will be to slightly increase the SNR (by a few) required for a confident detection of anisotropy.

B. Quadrupole injections

The nature of a quadrupole anisotropy in a GW-source distribution, with the maximal quadrupole angular distribution allowed under the constraint (41), which is $\propto 3\cos^2\theta$, can be seen in Fig. 13(a). Figure 14 shows the distinct angular correlations for all pulsars in the array when there is a quadrupole anisotropy in the GW-source population. We see that a quadrupole anisotropy, which in this case manifests itself through the angular distribution of sources being $\propto 1 + (3\cos^2\theta - 1)$, induces a larger deviation of the correlations from the Hellings and Downs curve than the dipole. We might expect then that the assumption

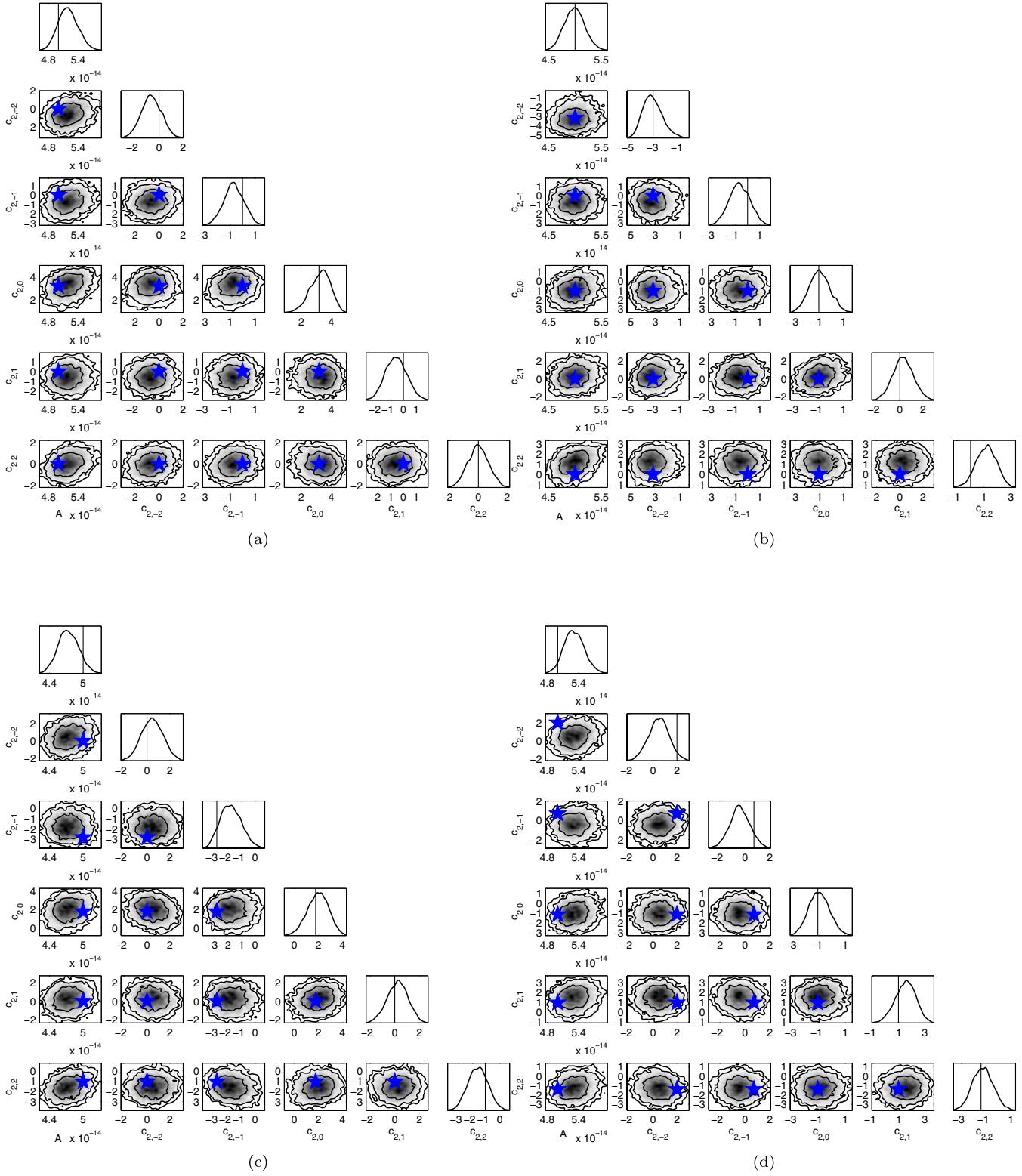


FIG. 16 (color online). The results of constraining the quadrupole-anisotropy coefficients, $c_{2,m}$, in data sets containing quadrupole anisotropies with a variety of configurations. The GW-source distributions corresponding to the injected anisotropies in each of these data sets are visualized in Fig. 13. The parameters of the injected anisotropies are (a) $\{c_{0,0} = 2\sqrt{\pi}; c_{2,0} = 4\sqrt{\pi}/5\}$, (b) $\{c_{0,0} = 2\sqrt{\pi}; c_{2,-2} = -0.9c_{0,0}; c_{2,0} = -0.3c_{0,0}\}$, (c) $\{c_{0,0} = 2\sqrt{\pi}; c_{2,-1} = -0.8c_{0,0}; c_{2,0} = 0.5c_{0,0}; c_{2,2} = -0.3c_{0,0}\}$, (d) $\{c_{0,0} = 2\sqrt{\pi}; c_{2,-2} = 0.56c_{0,0}; c_{2,-1} = 0.20c_{0,0}; c_{2,0} = -0.32c_{0,0}; c_{2,1} = 0.27c_{0,0}; c_{2,2} = -0.40c_{0,0}\}$.

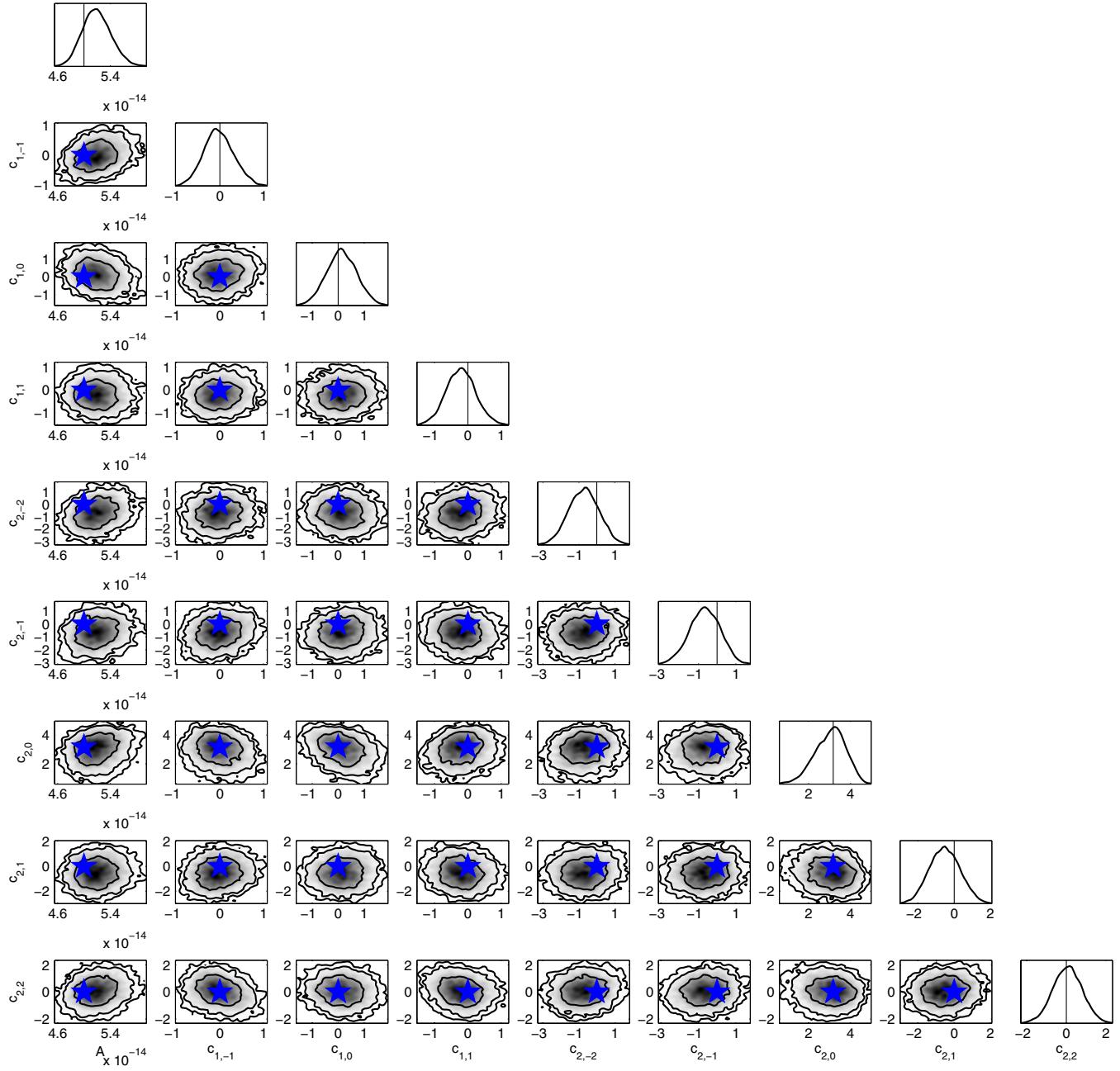


FIG. 17 (color online). A search over all anisotropy coefficients, c_{lm} , up to, and including, quadrupole order. The injected anisotropy was maximal quadrupole with $c_{0,0} = 2\sqrt{\pi}$ and $c_{2,0} = 4\sqrt{\pi/5}$, and all other dipole/quadrupole coefficients equal to zero. As can be seen, the anisotropic-search pipeline fully recovered all details of the injected anisotropy, as well as the amplitude of the GWB characteristic strain spectrum, A .

of isotropy will bias our results to a greater extent than in the dipole-anisotropy case.

We perform a similar analysis to Fig. 9, generating many data sets with fixed uncorrelated white-noise levels of 100 ns, but varying the amplitude of the GWB to simulate different SNRs [see Eq. (43)]. As in the dipole case, we vary the amplitude of the GWB between 5×10^{-16} – 5×10^{-14} , giving an effective SNR span of 0.1–10 and at each SNR we generate 10 data set realizations. The only

parameter we search for in the analysis of each data set is the amplitude, A , of the background. Hence, as before, this study represents a best-case assessment of when we will be able to identify the presence of a quadrupole anisotropy [of the form $P(\theta, \phi) = 3\cos^2\theta$] in the data. In reality, we will not know the direction of the quadrupole anisotropy in advance, which introduces extra dimensionality into the parameter space that must be searched over. This will tend to reduce the log-evidence value and so larger SNRs will be

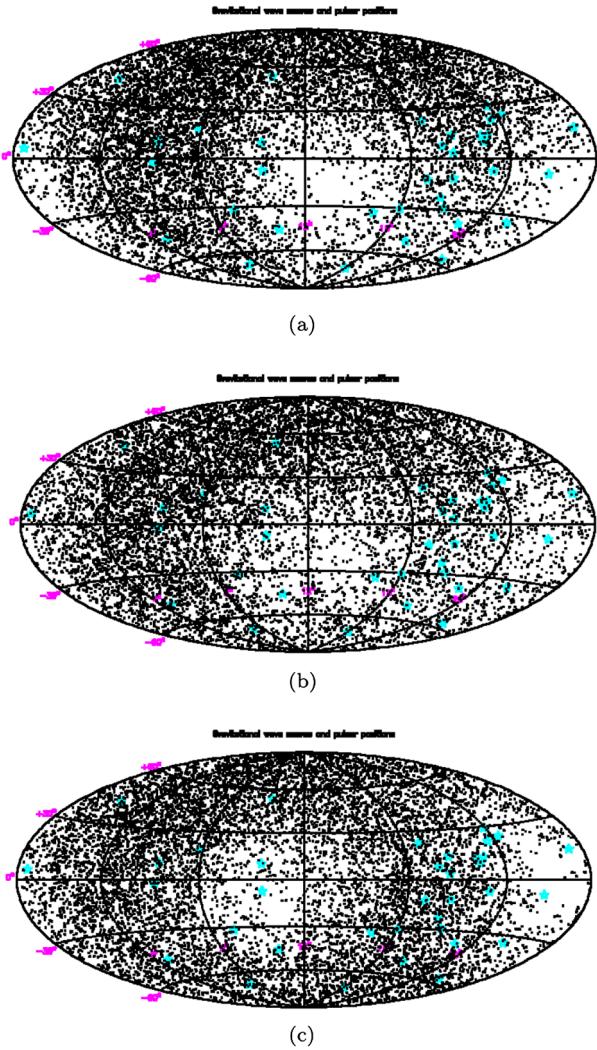


FIG. 18 (color online). Skymap screenshots from the GWanisobkrd TEMPO2 plugin showing the angular distribution of GW sources constituting various GWB realizations. The title of each plot is “Gravitational wave sources and pulsar positions,” with pulsars shown as light-blue stars and GW sources shown as black points. Lines of declination and right ascension are $[0^\circ, \pm 30^\circ, \pm 60^\circ]$ and $[4h, 8h, 12h, 16h, 20h]$, respectively. In (a) we show the skymap for the distribution of GW sources implied by a GWB with anisotropy coefficients $\{c_{0,0} = 2\sqrt{\pi}; c_{1,-1} = -0.4c_{0,0}; c_{1,0} = 0.4c_{0,0}; c_{2,0} = 0.20c_{0,0}; c_{2,2} = -0.45c_{0,0}\}$. In (b) we plot the skymap implied by the maximum *a posteriori* anisotropic configuration from an analysis with our Bayesian search pipeline, where we have imposed a hard physical prior (discussed in text). Finally, (c) shows the skymap implied by the maximum-likelihood anisotropic configuration, where we see that, despite not having the extra constraining influence of the physical prior, we are still able to adequately recover the distribution of GW sources.

required to identify quadrupole anisotropies in practice. As in the dipole case, one should note that because the injected noise is white-uncorrelated, and we have compressed to preserve information about a steep red-noise process, even at SNR = 0.1 the GWB is loud enough to be detected with

substantial evidence. At this SNR the largest noise-versus-(isotropic-)GWB log-Bayes factor is 0.3, with a mean and median of -9.3 and -5.8 , respectively. As discussed previously, at SNR = 0.1 the GWB is loud enough such that we can detect a background which has *approximate* Hellings and Downs form, but not yet loud enough to allow inference of anisotropy-induced deviations.

The results of this study can be seen in Fig. 15. These results are quite similar to the dipole case—the presence of anisotropy becomes identifiable at a SNR of about 1 and the evidence is decisive for a SNR of ~ 5 . We provide a calibration point at SNR = 10, where we search over $\{A, c_{2,m}\}$ for the 10 data set realizations at this SNR. As expected, we see that the increased dimensionality of the search penalizes the recovered Bayes factors; however the average log-Bayes factor implies that the presence of anisotropy may become visible with substantial evidence at SNR = 10, but we will require SNR > 10 for it to be clearly visible.

As in the dipole case, we test the robustness of the generalized anisotropic formalism by generating data sets with quadrupole anisotropies in a variety of configurations. We begin with the power contained in one quadrupole moment, then allow all the c_{lm} 's to be nonzero. The GW-source distribution for each of the tested configurations is shown in Fig. 13. As can be seen in Fig. 16, the generalized anisotropic-search pipeline recovers anisotropy coefficients which are consistent with the injected values. However the dimensionality is significantly larger than an isotropic search, and even in these high SNR data sets the posterior probability distributions are consistent with a wide range of possible anisotropies. This is partly due to the final three data sets having submaximal quadrupole anisotropy.

Finally, we repeat the analysis of the maximal-quadrupole data set [with $P(\theta, \phi) = 3\cos^2\theta$], but relax the assumption of quadrupole anisotropy in the analysis. The results of this analysis are shown in Fig. 17, where we have accurately recovered the nature of the injected anisotropy. However, as observed in the case where we performed an arbitrary-anisotropy search on a pure-dipole data set, the increased dimensionality of the arbitrary-anisotropy search widens the recovered posterior distributions to be consistent with a wide range of anisotropies.

C. Arbitrary anisotropy

1. Arbitrary-anisotropy search, and a physical prior

We now test our search pipeline on a data set with an arbitrary anisotropy spread among monopole, dipole and quadrupole components. The anisotropy we have chosen has a positive probability distribution for the placement of GW sources on the sky. The coefficients characterizing the anisotropy are $\{c_{0,0} = 2\sqrt{\pi}; c_{1,-1} = -0.4c_{0,0}; c_{1,0} = 0.4c_{0,0}; c_{2,0} = 0.20c_{0,0}; c_{2,2} = -0.45c_{0,0}\}$, with all other anisotropy coefficients set to zero. The skymap for the

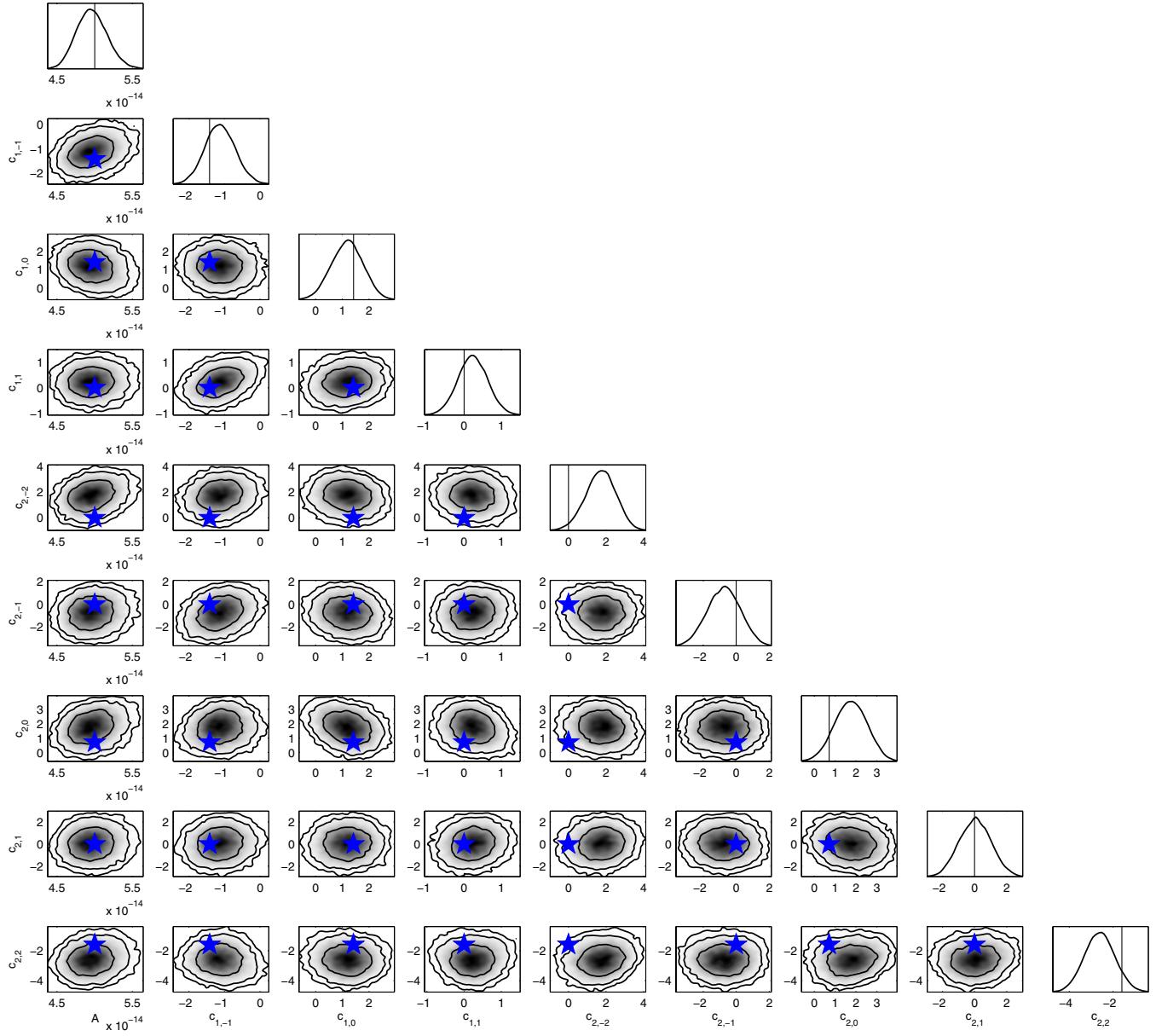


FIG. 19 (color online). A full anisotropic search on an arbitrary-anisotropy data set. The details of the data set are $\{c_{0,0} = 2\sqrt{\pi}; c_{1,-1} = -0.4c_{0,0}; c_{1,0} = 0.4c_{0,0}; c_{2,0} = 0.20c_{0,0}; c_{2,2} = -0.45c_{0,0}\}$. The prior is wide and flat, where $c_{l,m} \in [-5, 5]$, such that all anisotropy-coefficient combinations are *a priori* equally likely.

distribution of GW sources on the sky corresponding to these anisotropy coefficients is shown in Fig. 18(a).

Taking a wide flat prior on the anisotropy coefficients as before, the results of a search are shown in Fig. 19. The injected anisotropy coefficients all lie within the 2σ contours of the posterior distributions for the recovered coefficients. However, as commented on when the pure dipole and quadrupole data sets were subjected to an arbitrary-anisotropy search, the dimensionality that degrades the precision with the coefficients can be recovered, so that the posterior is consistent with a wide range of anisotropic distributions.

We now address this with a more physically motivated prior choice. Until now we have assumed that all

combinations of anisotropy coefficients are *a priori* equally likely. However, when generating these anisotropic data sets we have had to take into account that some combinations do not represent physical anisotropies, since they fail the condition [Eq. (41)] that the PDF for the distribution of GW sources on the sky is positive at all sky locations,

$$P(\hat{\Omega}) \propto \frac{dN}{d\hat{\Omega}} \propto \sum_{l,m} c_{lm} Y_{lm}(\hat{\Omega}) \geq 0, \quad \forall \hat{\Omega}. \quad (45)$$

Therefore, if we are to correctly perform parameter estimation on anisotropic data sets, this condition is the

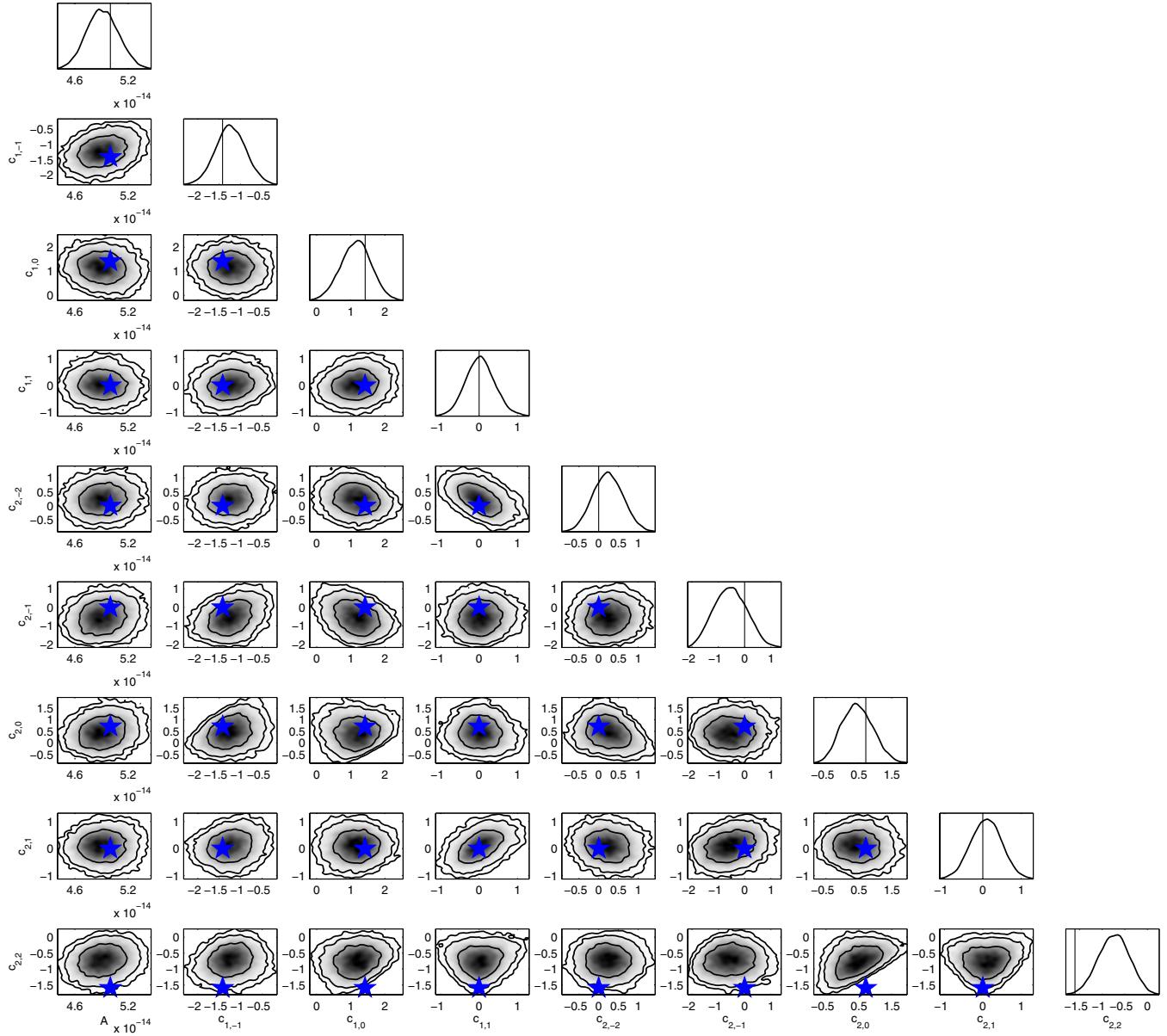


FIG. 20 (color online). A full anisotropic search on the same arbitrary-anisotropy data set as in Fig. 19. The anisotropy coefficients used to generate the data set are $\{c_{0,0} = 2\sqrt{\pi}; c_{1,-1} = -0.4c_{0,0}; c_{1,0} = 0.4c_{0,0}; c_{2,0} = 0.20c_{0,0}; c_{2,2} = -0.45c_{0,0}\}$. The prior is the only physically meaningful prior on possible anisotropic distributions. This corresponds to only sampling in regions where $\sum_{l,m} c_{lm} Y_{lm}(\hat{\Omega}) \geq 0, \forall \hat{\Omega}$.

only physically meaningful prior on the anisotropy coefficients. We impose it by setting the likelihood of any combination of $c_{l,m}$ which fails this condition to be very low, so as to prevent sampling in these regions. In practice this is achieved by pixelating the sky with a 20×20 grid in $[\cos \theta, \phi]$, and testing the condition in each pixel. If the condition is not met in every pixel, then the sample is rejected. The result of such an application of this prior on the previously analyzed arbitrary-anisotropy data set is shown in Fig. 20. For the most part, the recovered posterior distributions are consistent with the injected anisotropy. The exception is for the $c_{2,2}$

coefficient, where the injected amplitude lies in the tails of the recovered posterior distribution. The reason for this apparent inconsistency is actually quite straightforward. If we attempt to generate a data set with all anisotropy coefficients as before, except for $c_{2,2}$, which we lower from $-0.45c_{0,0}$ to $-0.47c_{0,0}$, then the GWanisobkrd plugin fails, since this PDF for the placement of GW sources on the sky is negative in some sky locations. Thus the reason that the injected $c_{2,2}$ amplitude is in the tails of the recovered posterior distribution is because this value is *a priori* disfavored, and very close to the boundary of nonphysicality.

TABLE III. Testing the ability of the anisotropic-search pipeline to identify the nature of the injected anisotropy. We subject data sets with different anisotropies (isotropic, dipole, quadrupole, arbitrary) to isotropic, dipole, quadrupole and arbitrary searches. We are interested in evidence values, and so we adopt wide flat priors on the anisotropy coefficients ($c_{l,m} \in [-5, 5]$) since the physical prior discussed in the previous subsection is actually very informative, truncating regions of high likelihood. The details of the injected arbitrary-anisotropy data set are $\{c_{0,0} = 2\sqrt{\pi}; c_{1,-1} = -0.4c_{0,0}; c_{1,0} = 0.4c_{0,0}; c_{2,0} = 0.20c_{0,0}; c_{2,2} = -0.45c_{0,0}\}$. We quote log-Bayes factors for the anisotropic models versus the isotropic model for each combination of injected anisotropy and model used to test for anisotropy.

		Injected anisotropy, {details}			
		$l = 0$ $\{c_{0,0} = 2\sqrt{\pi}\}$	$l = 0, 1$ $\{c_{0,0} = 2\sqrt{\pi}, c_{1,0} = 2\sqrt{\pi}/3\}$	$l = 0, 2$ $\{c_{0,0} = 2\sqrt{\pi}, c_{2,0} = 4\sqrt{\pi}/5\}$	$l = 0, 1, 2$ $\{\text{details in caption}\}$
Tested anisotropy, {search parameters}	$l = 0, \{A\}$	0.0	0.0	0.0	0.0
	$l = 0, 1, \{A, c_{1,m}\}$	-5.8	2.2	-2.7	5.8
	$l = 0, 2, \{A, c_{2,m}\}$	-5.6	-4.4	2.8	8.9
	$l = 0, 1, 2, \{A, c_{l,m}\}$	-10.4	-3.9	-2.3	9.7
	True, $\{A\}$	0.0	8.2	7.7	17.6

Despite the apparent inconsistency of the $c_{2,2}$ coefficient, we are still able to accurately recover the distribution of GW sources on the sky. In Fig. 18 we compare the skymaps of GW-background realizations for the injected anisotropy, a maximum *a posteriori* anisotropy of an analysis with the physical prior imposed on the $c_{l,m}$ coefficients, and the maximum-likelihood anisotropy where the relative power in $l = 0$ has been boosted to ensure physicality over the entire sky. As can be seen, the anisotropic-search pipeline has correctly recovered the features of this background. Even the maximum-likelihood anisotropic configuration, despite being unphysical, can have the relative power of its isotropic component boosted to ensure physicality over the entire sky, producing a GW source distribution which is adequately close to the injected distribution.

2. Anisotropy misfits

We now test the ability of the anisotropic-search pipeline to recover the *nature* of the GW-background anisotropy. Taking isolated $\text{SNR} = 10$ realizations of isotropic, dipole, quadrupole and arbitrary anisotropy data sets, we computed the evidence for isotropic, dipole, quadrupole and arbitrary models. We also computed the evidence for the true, injected anisotropy, searching only over the amplitude of the background. The results are shown in Table III, where all numbers indicate log-Bayes factors relative to an isotropic model. For these evidence calculations, we adopt the flat priors on the anisotropy coefficients. Although in practice we should be imposing the physicality conditions, we have found that the physical prior studied in the previous subsection is actually highly informative, and can truncate regions of high likelihood. The anisotropic configurations we have studied are close to maximal, which means the maximum-likelihood parameter regions will be close to nonphysicality. Taking the flat priors can therefore help improve the anisotropic evidence, since the

high likelihood region stretches beyond the boundary imposed by physicality. Thus, for these cases, the unphysical evidence seems to be a better statistic to use to distinguish models than the physical evidence. Further investigations will be required to determine if this is true more generally, in particular for cases in which the anisotropy is weak. Interpreting the unphysical distributions on the coefficients is more difficult, since they give rise to negative probabilities for certain regions of the sky. However, increasing the relative power of the isotropic $l = 0$ mode can always be used to make the distribution physical, and provide source distributions which are adequately close to the injected distribution. The effect of the unphysical anisotropic coefficients, c_{lm} , on our search is to allow the GWB correlations to deviate by more from the Hellings and Downs curve, allowing some of the noise to be absorbed into these coefficients. This seems to help for the limited number of cases we have considered but further work is required to understand which prior works better in more general situations.

Referring to Table III, the isotropic model is highly favored when describing the isotropic data set. The true dipole anisotropy is substantially favored over an isotropic model when describing the dipole data set, and likewise the quadrupole anisotropy is substantially favored over an isotropic model when describing the quadrupole data set. Finally, in the model comparison of the arbitrary-anisotropy data set, any form of anisotropy is favored over the isotropic model. However, an arbitrary-anisotropic search is favored over both the dipole and quadrupole searches, but in the latter case the degree to which the arbitrary-anisotropic search is favored is inconclusive. We suspect the reason for this can be understood in terms of how the different anisotropies affect the scatter of correlations around the isotropic Hellings and Downs curve, particularly for pulsar angular separations where there are many pulsar pairings. This is shown in Fig. 21, where we see that in the region where there are many

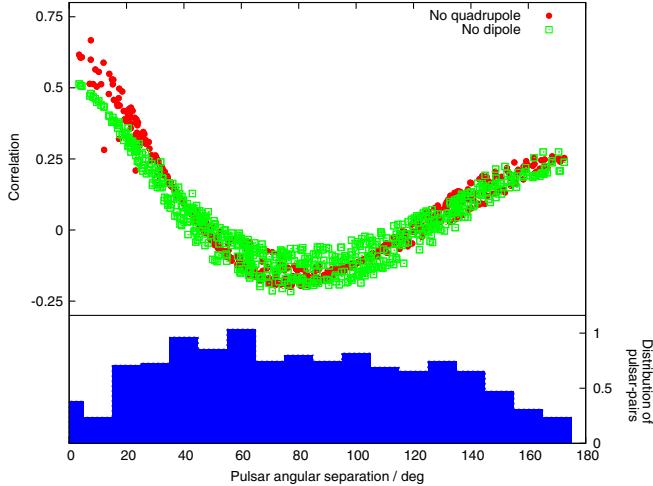


FIG. 21 (color online). We show the angular correlations for pulsar pairings with nonzero angular separation in an arbitrary-anisotropy data set, for the cases where we ignore the contribution from the dipole anisotropy and when we ignore the quadrupole anisotropy. In the region of pulsar angular separation where there are many distinct pulsar pairings, the quadrupole-only anisotropy dominates the deviation of correlations from the isotropic Hellings and Downs curve. This may explain why a quadrupole-only model is almost as good as an arbitrary-anisotropy model when describing the arbitrary-anisotropy data set. Note that only nonzero angular-separation pairings are shown in the angular-correlation subplot, whereas all pairings are included in the pulsar-pairing histogram.

distinct pulsar pairings the quadrupole anisotropy induces a larger scatter and deviation from the isotropic Hellings and Downs curve than the dipole anisotropy. Hence, in this case, a quadrupole-only model is almost as good as an arbitrary-anisotropy model when describing the arbitrary-anisotropy data set.

VIII. CONCLUSIONS

We have systematically investigated the robustness and veracity of a Bayesian search pipeline capable of probing anisotropic gravitational-wave backgrounds using pulsar timing arrays. If we have some form of continuous anisotropy in the energy density of such a GWB, or anisotropy through sparseness of sources comprising our background, then the assumption of isotropy may no longer be an appropriate model with which to perform searches. Recent work by Cornish and Sesana [41], and Mingarelli *et al.* [42] has focused on developing new formalisms and techniques with which to generalize the current search methodologies, which all fall within the categories of either stochastic isotropic searches or single-source searches.

We employed the formalism of Mingarelli *et al.* [42], which generalizes the cross correlation of GWB-induced TOA deviations between pulsars to anisotropic backgrounds. In an isotropic GWB this cross correlation has a

unique, distinctive signature, being a function only of the angular separation between pulsars on the sky. It is commonly referred to as the Hellings and Downs curve [40]. In the more generalized formalism, the distribution of GWB power is decomposed as a function of spherical harmonics, and a set of correlation basis functions uniquely defined for a particular PTA configuration, although the aforementioned authors have identified useful analytic expressions for these basis functions up to quadrupole order. Unsurprisingly, the monopole solution, describing the induced correlation for the isotropic mode, is the usual Hellings and Downs curve. Crucially, the correlation basis functions for dipole anisotropies and above are no longer independent of the absolute sky location of the pulsar, being strongly dependent on the placement of the pulsars relative to the distribution of GWB power.

To thoroughly investigate this formalism, we have developed a suite of new plugins for the popular pulsar timing analysis software package, TEMPO2. This permits the user to have complete control over the generation of realistic-format data sets, including specifying the array of pulsars, the various noise processes affecting these pulsars, and the ability to inject a GWB signal into these simulated data sets. We have generalized this to inject continuous anisotropies into the GWB signals, and also to permit the inclusion of non-Einsteinian polarization modes in the GW signals.

By generalizing the cross correlation of pulsar timing residuals in a time-domain Bayesian search, we have probed the level of anisotropy in the distribution of GW source populations making up many background realizations. In particular, we have found that our pipeline can infer the direction of a strong dipole anisotropy in a loud GWB. Not only that, but by performing Bayesian evidence evaluation using the MULTINEST package [56,57], such a strong dipole anisotropy can be favored over the standard isotropic search, where we see that the latter is no longer an appropriate model to use. By evaluating the evidence of data sets with the anisotropy fixed at the injected levels, and subsequently comparing to the isotropic evidence, we find that the best-case scenario for having decisive evidence for a strong-dipole anisotropy occurs when the induced rms residual of the GWB is approximately five times greater than the noise rms residual. When a more general search over the dipole anisotropy coefficients is performed, the log-Bayes factors are degraded, but substantial evidence for the anisotropic signal may become visible at $\text{SNRs} \geq 10$. We performed a similar analysis with uncorrelated red noise affecting the pulsars, finding a similar level of required GWB loudness in order to begin having convincing evidence of anisotropy.

We extended our search to data sets with injected GWBs containing quadrupole anisotropies in a variety of different configurations. Again, we found that the search pipeline was able to recover posterior probability distributions for

the anisotropic-decomposition coefficients, finding agreement with injected values. We repeated our analysis of the best-case scenario for when we could have convincing evidence of anisotropy in the GWB, this time for the quadrupole anisotropy. As before, in this absolute best-case scenario, where we fix the tested anisotropy to the injected values, we still require the rms residual induced by the GWB to be approximately five times greater than the noise rms residual. When a more general search over the quadrupole anisotropy coefficients is performed, the log-Bayes factors are again degraded, but substantial evidence for the anisotropic signal is possible at $\text{SNR} = 10$, such that clear evidence would require $\text{SNRs} \geq 10$. Further work is needed here, in order to systematically evaluate the Bayesian evidence at different GWB loudnesses without suffering from the curse of high dimensionality, which can have a detrimental effect when comparing to isotropic models.

Finally, we tested our search pipeline on a data set with the distribution of GWB power spread over monopole, dipole and quadrupole. While the recovered maximum-likelihood anisotropy produces a distribution of GW sources which adequately mimics the injected distribution, the recovered probability distribution of anisotropy coefficients is rather wide since we are now searching over nine dimensions. Thus, we could conceivably be consistent with a wide range of anisotropic configurations. Hence we propose a hard prior on the anisotropy coefficients, corresponding to the physicality of the implied probability distribution for the placement of sources on the sky. If this PDF happens to be negative anywhere on the sky for a proposed set of anisotropy coefficients, then this point is discarded in the sampling process. This is actually quite an informative prior, such that if we are studying a data set with an anisotropy close to nonphysicality, then the prior may actually truncate regions of high likelihood in parameter space. This could conceivably have a negative impact on evidence evaluation; however it is a very useful tool in

parameter estimation. The maximum *a posteriori* anisotropy when this physicality prior is imposed shows a greater match to the injected anisotropy. We intend to study the effects of this prior in future work.

The search for angular structure in the GWB will be the obvious next step in GWB PTA searches, permitting us to investigate the clustering of sources or GW hotspots. This paper has robustly tested an anisotropic stochastic search pipeline on data sets with a continuous anisotropy in the angular distribution of sources constituting the injected background. We intend to follow up this first study with a further investigation of the consequences of imposing a hard physical prior, as well as applying this search pipeline to data sets with GWBs consisting of realistic populations of inspiraling SMBHBs. This latter investigation will allow us to understand the effects of source sparseness and background finiteness (which occur at higher GW frequencies as the stochasticity of the strain signal breaks down) on the ability to constrain the properties of a GWB. The next several years should see rapid advances in the field of PTA searches for GWs, as we move beyond the standard paradigm of an isotropic search; develop faster, more robust algorithms; and perform the first combined search using the efforts and resources from all three PTAs.

ACKNOWLEDGMENTS

We thank Chiara Mingarelli and Alberto Vecchio for many useful discussions and Alberto Sesana for supplying the data used to generate the figures in Sec. II. S. R. T. is supported by the STFC, and acknowledges the support of the Royal Astronomical Society. J. R. G. is supported by the Royal Society. This work was performed using the Darwin Supercomputer of the University of Cambridge High Performance Computing Service (<http://www.hpc.cam.ac.uk/>), provided by Dell Inc. using Strategic Research Infrastructure Funding from the Higher Education Funding Council for England.

-
- [1] G. M. Harry (LIGO Scientific Collaboration), *Classical Quantum Gravity* **27**, 084006 (2010).
 - [2] The Virgo Collaboration, Virgo Technical Report No. VIR-0027A-09, 2009.
 - [3] K. Somiya (KAGRA Collaboration), *Classical Quantum Gravity* **29**, 124007 (2012).
 - [4] J. Aasi, J. Abadie, B. P. Abbott, R. Abbott, T. D. Abbott, M. Abernathy, T. Accadia, F. Acernese *et al.* (LIGO Scientific and Virgo Collaborations), [arXiv:1304.0670](https://arxiv.org/abs/1304.0670).
 - [5] eLISA/NGO, <http://www.elisa-ngo.org/>.
 - [6] P. A. Seoane, S. Aoudia, H. Audley, G. Auger, S. Babak, J. Baker, E. Barausse, S. Barke *et al.* (T. e. Consortium), [arXiv:1305.5720](https://arxiv.org/abs/1305.5720).
 - [7] R. S. Foster and D. C. Backer, *Astrophys. J.* **361**, 300 (1990).
 - [8] S. T. McWilliams, J. P. Ostriker, and F. Pretorius, [arXiv:1211.4590](https://arxiv.org/abs/1211.4590).
 - [9] S. T. McWilliams, J. P. Ostriker, and F. Pretorius, [arXiv:1211.5377](https://arxiv.org/abs/1211.5377).
 - [10] A. Sesana, *Mon. Not. R. Astron. Soc.* **433**, L1 (2013).
 - [11] V. M. Kaspi, J. H. Taylor, and M. F. Ryba, *Astrophys. J.* **428**, 713 (1994).
 - [12] D. N. Matsakis, J. H. Taylor, and T. M. Eubanks, *Astron. Astrophys.* **326**, 924 (1997).
 - [13] M. V. Sazhin, *Sov. Astron.* **22**, 36 (1978).
 - [14] S. Detweiler, *Astrophys. J.* **234**, 1100 (1979).

- [15] F. B. Estabrook and H. D. Wahlquist, *Gen. Relativ. Gravit.* **6**, 439 (1975).
- [16] W. L. Burke, *Astrophys. J.* **196**, 329 (1975).
- [17] M. C. Begelman, R. D. Blandford, and M. J. Rees, *Nature (London)* **287**, 307 (1980).
- [18] E. S. Phinney, [arXiv:astro-ph/0108028](https://arxiv.org/abs/astro-ph/0108028).
- [19] A. H. Jaffe and D. C. Backer, *Astrophys. J.* **583**, 616 (2003).
- [20] J. S. B. Wyithe and A. Loeb, *Astrophys. J.* **590**, 691 (2003).
- [21] A. Sesana, A. Vecchio, and C. N. Colacino, *Mon. Not. R. Astron. Soc.* **390**, 192 (2008).
- [22] L. P. Grishchuk, *Pis'ma Zh. Eksp. Teor. Fiz.* **23**, 326 (1976).
- [23] L. P. Grishchuk, *Phys. Usp.* **48**, 1235 (2005).
- [24] W. Zhao, *Phys. Rev. D* **83**, 104021 (2011).
- [25] W. Zhao, Y. Zhang, X.-P. You, and Z.-H. Zhu, *Phys. Rev. D* **87**, 124012 (2013).
- [26] A. Vilenkin, *Phys. Rev. D* **24**, 2082 (1981).
- [27] A. Vilenkin, *Phys. Lett.* **107B**, 47 (1981).
- [28] S. Ölmez, V. Mandic, and X. Siemens, *Phys. Rev. D* **81**, 104028 (2010).
- [29] S. A. Sanidas, R. A. Battye, and B. W. Stappers, *Phys. Rev. D* **85**, 122003 (2012).
- [30] T. Damour and A. Vilenkin, *Phys. Rev. D* **71**, 063510 (2005).
- [31] F. A. Jenet, G. B. Hobbs, W. van Straten, R. N. Manchester, M. Bailes, J. P. W. Verbiest, R. T. Edwards, A. W. Hotan, J. M. Sarkissian, and S. M. Ord, *Astrophys. J.* **653**, 1571 (2006).
- [32] V. Ravi, J. S. B. Wyithe, G. Hobbs, R. M. Shannon, R. N. Manchester, D. R. B. Yardley, and M. J. Keith, *Astrophys. J.* **761**, 84 (2012).
- [33] P. B. Demorest, R. D. Ferdman, M. E. Gonzalez, D. Nice, S. Ransom, I. H. Stairs, Z. Arzoumanian, A. Brazier, S. Burke-Spoloar, S. J. Chamberlin *et al.*, *Astrophys. J.* **762**, 94 (2013).
- [34] D. R. B. Yardley, W. A. Coles, G. B. Hobbs, J. P. W. Verbiest, R. N. Manchester, W. van Straten, F. A. Jenet, M. Bailes, N. D. R. Bhat, S. Burke-Spoloar *et al.*, *Mon. Not. R. Astron. Soc.* **414**, 1777 (2011).
- [35] R. van Haasteren, Y. Levin, G. H. Janssen, K. Lazaridis, M. Kramer, B. W. Stappers, G. Desvignes, M. B. Purver, A. G. Lyne, R. D. Ferdman *et al.*, *Mon. Not. R. Astron. Soc.* **414**, 3117 (2011).
- [36] European Pulsar Timing Array, <http://www.epta.eu.org/>.
- [37] North American Nanohertz Observatory for Gravitational Waves, <http://nanograv.org/>.
- [38] Parkes Pulsar Timing Array, <http://www.atnf.csiro.au/research/pulsar/ppta/>.
- [39] International Pulsar Timing Array, <http://www.ippta4gw.org/>.
- [40] R. W. Hellings and G. S. Downs, *Astrophys. J.* **265**, L39 (1983).
- [41] N. J. Cornish and A. Sesana, [arXiv:1305.0326](https://arxiv.org/abs/1305.0326).
- [42] C. M. F. Mingarelli, T. Sidery, I. Mandel, and A. Vecchio, [arXiv:1306.5394](https://arxiv.org/abs/1306.5394) [Phys. Rev. D (to be published)].
- [43] G. B. Hobbs, R. T. Edwards, and R. N. Manchester, *Mon. Not. R. Astron. Soc.* **369**, 655 (2006).
- [44] R. T. Edwards, G. B. Hobbs, and R. N. Manchester, *Mon. Not. R. Astron. Soc.* **372**, 1549 (2006).
- [45] G. Hobbs, F. Jenet, K. J. Lee, J. P. W. Verbiest, D. Yardley, R. Manchester, A. Lommen, W. Coles, R. Edwards, and C. Shettigara, *Mon. Not. R. Astron. Soc.* **394**, 1945 (2009).
- [46] M. Anholm, S. Ballmer, J. D. E. Creighton, L. R. Price, and X. Siemens, *Phys. Rev. D* **79**, 084030 (2009).
- [47] L. G. Book and É. É. Flanagan, *Phys. Rev. D* **83**, 024024 (2011).
- [48] K. Lee, F. A. Jenet, R. H. Price, N. Wex, and M. Kramer, *Astrophys. J.* **722**, 1589 (2010).
- [49] E. Poisson, *Classical Quantum Gravity* **25**, 209002 (2008).
- [50] K. J. Lee, F. A. Jenet, and R. H. Price, *Astrophys. J.* **685**, 1304 (2008).
- [51] B. Allen and A. C. Ottewill, *Phys. Rev. D* **56**, 545 (1997).
- [52] H. Jeffreys, *Theory of Probability*, International Series of Monographs on Physics (Clarendon Press, Oxford, 1983).
- [53] R. van Haasteren and Y. Levin, *Mon. Not. R. Astron. Soc.* **428**, 1147 (2013).
- [54] R. van Haasteren, *Mon. Not. R. Astron. Soc.* **429**, 55 (2013).
- [55] J. Skilling, in *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, edited by T. U. V. Fischer and R. Preuss (American Institute of Physics, New York, 2004), Vol. 735, p. 395.
- [56] F. Feroz and M. P. Hobson, *Mon. Not. R. Astron. Soc.* **384**, 449 (2008).
- [57] F. Feroz, M. P. Hobson, and M. Bridges, *Mon. Not. R. Astron. Soc.* **398**, 1601 (2009).
- [58] K. J. Lee, C. G. Bassa, G. H. Janssen, R. Karuppusamy, M. Kramer, R. Smits, and B. W. Stappers, *Mon. Not. R. Astron. Soc.* **423**, 2642 (2012).