

Establishing the significance of continuous gravitational-wave detections from known pulsars

Maximiliano Isi^{1,*}, Simone Mastrogiovanni^{2,†}, Matthew Pitkin³, and Ornella Juliana Piccinni⁴

¹*LIGO Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

²*AstroParticule et Cosmologie (APC), 10 Rue Alice Domon et Léonie Duquet, Université de Paris, CNRS, AstroParticule et Cosmologie, F-75013 Paris, France*

³*Department of Physics, Lancaster University, Lancaster LA1 4YB, United Kingdom*

⁴*INFN, Sezione di Roma, I-00185 Roma, Italy*



(Received 29 October 2020; accepted 2 December 2020; published 28 December 2020)

We present a method for assigning a statistical significance to detection candidates in targeted searches for continuous gravitational waves from known pulsars, without assuming the detector noise is Gaussian and stationary. We take advantage of the expected Doppler phase modulation of the signal induced by Earth's orbital motion, as well as the amplitude modulation induced by Earth's spin, to effectively blind the search to real astrophysical signals from a given location in the sky. We use this “sky shifting” to produce a large number of noise-only data realizations to empirically estimate the background of a search and assign detection significances, in a similar fashion to the use of time slides in searches for compact binaries. We demonstrate the potential of this approach by means of simulated signals, as well as hardware injections into real detector data. In a study of simulated signals in non-Gaussian noise, we find that our method outperforms another common strategy for evaluating detection significance. We thus demonstrate that this and similar techniques have the potential to enable a first confident detection of continuous gravitational waves.

DOI: [10.1103/PhysRevD.102.123027](https://doi.org/10.1103/PhysRevD.102.123027)

I. INTRODUCTION

In addition to short-lived gravitational waves (GWs) from compact-binary coalescences like those observed so far [1–8], ground-based detectors like the Advanced Laser Interferometer Gravitational-Wave Observatory (aLIGO) [9] and Virgo [10] are also expected to detect persistent, almost-monochromatic signals [11–17]. The primary potential source of such “continuous waves” (CWs) is rapidly spinning neutron stars with an asymmetry in their moment of inertia [18]. This includes galactic pulsars known from electromagnetic observations, which are a main target for searches for continuous signals in LIGO and Virgo data [16,19,20]. The detection of gravitational waves from any of these sources would provide a new wealth of astrophysical information, as well as invaluable opportunities to learn about fundamental physics (see, e.g., [21] for a recent review).

There exist a number of efforts to detect gravitational waves from known pulsars [15,16,22,23]. However, an outstanding problem affecting all of these searches is the lack of a well-defined procedure to establish the statistical significance of potential detections without making the

assumption that the instrumental noise is Gaussian and (semi-)stationary. Consequently, if evidence for a continuous wave from a known pulsar was found today, we would be unable to establish, with certainty, the probability for this to have arisen from a spurious noise artifact. The need for a systematic and robust way of computing detection significance in the presence of non-Gaussian noise has already become apparent with the appearance of hard-to-diagnose outliers in recent searches in actual aLIGO data [16,24].

Establishing a robust procedure to assign significance is challenging because the noise artifacts that limit the searches are intrinsically unpredictable and cannot be modeled from first principles. Given this, we may instead attempt to empirically determine the response of the different searches to real detector noise in the *absence* of astrophysical signals. Armed with such knowledge, we would then be able to analyze actual data, or “foreground,” and produce empirical likelihood ratios (or other measures of detection confidence, like p -values) for the presence of an astrophysical signal vs just instrumental noise, Gaussian or otherwise. This requires several instances of “background”—that is, instrumental data that are known to contain no astrophysical signals, while still retaining all statistical properties representative of real instrumental noise.

*maxisi@mit.edu

†simone.mastrogiovanni@apc.in2p3.fr

Ideally, one would obtain background distributions by physically isolating the instruments from the environment to shield them from actual signals. Because this is impossible in the case of gravitational waves, we must attempt to replicate this shielding digitally after the data have been recorded. Several techniques exist to do this when looking for gravitational-wave transients, the most straightforward of which is probably the use of “time slides”: The outputs of different detectors are shifted relative to each other by time offsets longer than the light-travel time between them [25,26]. This ensures the spuriousness of any signal candidate left in the multidetector data thus produced, hence allowing us to estimate how likely it is for noise to mimic a signal.

The direct analog of time slides in the context of continuous waves would be “frequency slides”: a misalignment of the frequency-domain data of different detectors. However, our ability to effect such frequency shifts is limited by the frequency resolution of the searches (of the order of inverse observation time), and the fact that the properties of actual instrumental noise are heavily dependent on frequency—not only due to a frequency-dependent power spectral density, but also to varying populations of narrow-band noise features. By the same token, time slides themselves would not be feasible in transient analyses if the noise properties of the detectors changed rapidly compared to the duration of a typical signal.

In light of this, here we propose a simple method for estimating the background of searches for continuous gravitational waves by analyzing data assuming an incorrect sky location for the targeted source. This “sky shifting” takes advantage of the expected Doppler modulation of the signal due to the relative motion of the detector and source to effectively blind the search to real astrophysical signals. Similar methods have been developed for pulsar timing arrays [27]. We can use this to produce a large number of noise-only instantiations of data, so as to empirically estimate the background of a search and assign detection significances in the presence of actual detector noise. We demonstrate that this method can outperform another common strategy for estimating the background in realistic situations.

We begin by providing relevant background about continuous waves and targeted searches in Sec. II. We then introduce the sky-shifting method and explore its applicability in Sec. III. We demonstrate the efficacy of the strategy in Sec. IV with the aid of several examples based on both synthetic and actual detector noise. We conclude in Sec. V.

II. BACKGROUND

In this section, we review the basic morphology of continuous gravitational waves as measured by differential-arm detectors, with an emphasis on the timing corrections on which we will rely for sky shifting (Sec. II A 1). We also make a special point of discussing the relation between the

frequency resolution at which a signal is sampled and the ability to localize the source in the sky (Sec. II A 2). We next describe the key properties of noise in existing ground-based instruments as it pertains to searches for persistent signals (Sec. II B). Finally, we provide an overview of three staple search methods for these signals in LIGO and Virgo data (Sec. II C): the Bayesian time-domain method, and the frequentist five-vector and \mathcal{F} -statistic methods.

A. Continuous waves

1. Morphology

Continuous waves are nearly monochromatic gravitational perturbations with constant intrinsic amplitude that are expected to be sourced by some rapidly spinning bodies, like neutron stars. Within the context of standard physics, there are several ways in which a neutron star could emit CWs, but the most favored is the presence of a nonaxisymmetry in the star’s moment of inertia [28]. For this type of *triaxial*, nonprecessing source, such a GW will induce a strain in a differential-arm (quadrupolar) detector, like LIGO or Virgo, which can be written as

$$h(t) = h_0 \frac{1}{2} (1 + \cos^2 \iota) F_+(t; \psi) \cos \phi(t) + h_0 \cos \iota F_\times(t; \psi) \sin \phi(t), \quad (1)$$

where the $F_+(t; \psi)$ and $F_\times(t; \psi)$ factors respectively give the instrument’s response to the plus (+) and cross (×) GW polarizations, ι is the inclination angle between the spin axis of the source and the observer’s line of sight, $\phi(t)$ is the phase of the signal, and h_0 is an overall amplitude related to the properties of the source by

$$h_0 = \frac{16\pi^2 G}{c^4} \frac{\epsilon I_{zz} f_{\text{rot}}^2}{r}, \quad (2)$$

where r is the source distance, f_{rot} its rotation frequency around the principal axis z , \mathbf{I} the moment-of-inertia tensor, and $\epsilon \equiv (I_{xx} - I_{yy})/I_{zz}$ the equatorial ellipticity [18].

The antenna patterns $F_+(t; \psi)$ and $F_\times(t; \psi)$ encode the amplitude modulation of the signal due to the local geometric effect of a GW acting on a given detector. Thus, they implicitly depend on the relative location and orientation of source and detector by means of the source’s right-ascension α , declination δ , and polarization angle ψ . The latter gives the orientation of the frame in which the polarizations are defined, and we set it to be the angle between the line of nodes and the projection of the celestial north onto the plane of the sky. While α and δ are always well known, ψ generally is not, which is why we show this argument explicitly. Importantly, the antenna patterns acquire their time dependence from the rotation of Earth on its axis, and consequently have a characteristic period of a sidereal day ($\sim 10^{-5}$ Hz).

For a simple triaxial source, the GW frequency f is twice the rotational value f_{rot} , so we can write

$$\phi(t) = 2\phi_{\text{rot}}(t) + \phi_0, \quad (3)$$

where ϕ_{rot} is the rotational phase as measured via electromagnetic (EM) observations and ϕ_0 is a fiducial phase offset. The rotational frequency itself is almost constant, with a small spin-down due to energy loss into the environment (via GWs and other mechanisms), which means that the phase evolution can be well described by a simple Taylor expansion on τ , the time measured by a clock inertial with respect to the source:

$$\phi(t) = 2\pi \sum_{j=0}^N \frac{\partial_t^{(j)} f_0}{(j+1)!} [\tau(t) - T_0]^{(j+1)}. \quad (4)$$

Here $\partial_t^{(j)} f_0$ is the j th time derivative of the GW frequency measured at the fiducial time T_0 , and N is the order of the series expansion (1 or 2 suffices for most sources). Timing solutions are generally obtained through the pulsar timing package TEMPO2 [29]. These solutions have exquisite precision (frequency uncertainty of $10^{-12} \lesssim \delta f_{\text{rot}} \lesssim 10^{-8}$ Hz for most pulsars) and are the cornerstone of targeted searches for continuous waves from known pulsars.

The inertial time, τ in Eq. (3), is usually taken to be the time measured by a clock at the Solar System barycenter (SSB), which is itself assumed to be inertial with respect to the pulsar. In that case, τ can be written as a function of detector time t by taking into account some well-known, time-dependent offsets:

$$\tau(t) = t + \Delta_E(t) + \Delta_S(t) + \Delta_{\text{binary}}(t) + \Delta_R(t). \quad (5)$$

Here Δ_E is the Solar System Einstein delay, Δ_S is the Solar System Shapiro delay, Δ_{binary} is the delay originating from the motion of the pulsar in its binary (a term that vanishes for isolated sources) [30], and Δ_R is the kinematic delay due to the relative motion of the detector with respect to the source.

The timing correction of Eq. (5) is heavily dependent on the sky location of the targeted pulsar and will be the key to the sky-shifting method presented in Sec. III. The dependence on sky location is dominated by the last term in Eq. (5), Δ_R . This is sometimes known as the “Rømer delay” and encodes the Doppler modulation of the signal:

$$\Delta_R(t) = -\frac{\hat{\Omega} \cdot \vec{r}(t)}{c}, \quad (6)$$

where $\vec{r}(t)$ is a vector joining the SSB and the detector at any given time, $\hat{\Omega}$ is a unit vector pointing from the SSB in the direction of the source,¹ and c is the GW speed. For

¹The source-location vector $\hat{\Omega}$ can be treated as constant over the timescale of our observations.

practical purposes, \vec{r} is usually computed by first splitting it into three components:

$$\vec{r} = \vec{r}_\odot + \vec{r}_\oplus + \vec{R}, \quad (7)$$

with \vec{r}_\odot joining the SSB with the Sun’s center, \vec{r}_\oplus joining Sun and Earth, and \vec{R} going from Earth’s center to the detector on the surface. One can then use Solar System ephemerides, together with knowledge of the location of the detector on Earth and the source in the sky, to compute the Rømer correction at any given time.

The timing correction of Eq. (5) can be understood as inducing extrinsic frequency shifts to the signal, as seen by the detector. This is dominated by the Rømer term $\Delta_R(t)$, which results in a modulation at the frequency of Earth’s orbital rotation $\omega_{\text{orb}} \approx 2 \times 10^{-7}$ rad/s, as well as subdominant daily effects due to its spin $\omega_{\text{sid}} = 2\pi/(\text{sidereal day}) \approx 7 \times 10^{-5}$ rad/s. In the frequency domain, the effect of this correction is to spread the signal power across a narrow band centered on its intrinsic GW frequency, with a characteristic width of $\Delta f/f \approx \omega_{\text{orb}} r_\oplus / c \approx 10^{-4}$ (see e.g., [31]). This frequency modulation will be the key to our approach.

2. Frequency and sky resolution

The sky resolution is the minimum angular separation in the sky at which two, otherwise equal, sources could be distinguished. This is a function of the frequency resolution at which the signal is sampled, namely,

$$\delta f = 1/T, \quad (8)$$

for an observation time T . This frequency bin is related to a sky bin by the sky-location-dependent frequency modulation of Eq. (5). Thus, the angular resolution will be roughly given by the separation in the sky corresponding to a Rømer frequency shift of $\Delta f = \delta f$. In other words, we may define a bin around any point in the sky by the maximum angular distance one can move away from that point before the frequency shift caused by the modulation of Eq. (6) reaches a magnitude of $1/T$. Thus, the characteristic size of a bin defined this way will necessarily depend on the integration time.

Proceeding as above, we may cover the sky with a series of such bins to obtain a “sky grid” representing the resolvability of points in the sky as a function of angular location. Because the timing correction of Eq. (5) is dominated by Δ_R , which is itself mostly due to Earth’s orbital motion, such a sky grid will be most naturally defined in *ecliptic* coordinates to yield bin sizes given approximately by [32]

$$\delta\beta = \frac{1}{N_d |\sin\beta|}, \quad \delta\lambda = \frac{1}{N_d |\cos\beta|} \quad (9)$$

where β and λ are respectively the ecliptic latitude and longitude, and the scale factor (that represents the total number of Doppler bins) is

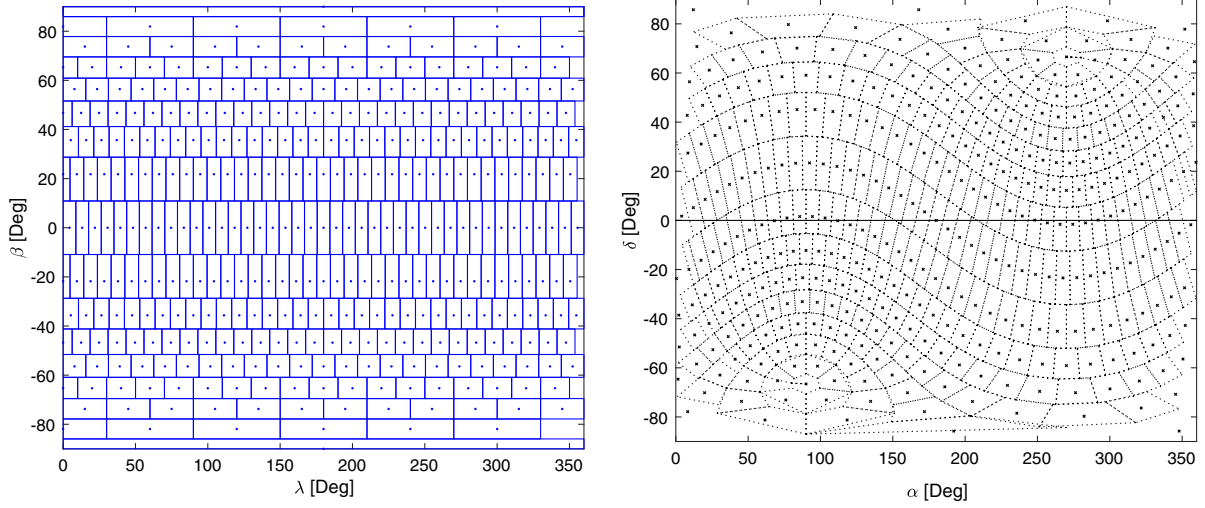


FIG. 1. Left: sky bins built using ecliptic coordinates for a search at a frequency of 60 Hz and an integration time of 1024 s. Right: the same sky grid for the same sky configuration in equatorial coordinates.

$$N_d = \Delta f \cdot f \cdot T = \frac{f \omega_{\text{orb}} r_{\oplus} T}{c} \quad (10)$$

for Earth’s orbital radius r_{\oplus} . As demonstrated in Fig. 1, the sky grid can be easily computed using ecliptic coordinates (left panel) and then rotated to equatorial coordinates (right panel). This is a conservative sky grid for an integration time of $T_{\text{coh}} = 1024\text{s}$ that implicitly assumes the power of the signal may be split over at most two frequency bins as a result of the timing correction—in practice, the characteristic size of the sky bins may be reduced, but the scaling with f and T will always be as in Eq. (10). In a full-coherent search that uses one year of data, the sky bins’ size is significantly reduced.

B. Detector noise

The output of ground-based gravitational-wave detectors is vastly dominated by instrumental noise [33,34]. For this reason, the weak continuous signals discussed in Sec. II A are expected to become visible only after long periods of coherently integrated observation. Understanding the statistical properties of the noise is critical to successfully detecting these signals.

For the most part, the noise in a given detector is well described as a Gaussian random process with a frequency-dependent (colored) power spectral density [33,34]. Gaussian noise has numerous convenient statistical properties that would drastically simplify many of LIGO and Virgo’s analyses. However, this idealization is far from perfect: The data are plagued with uncountable non-Gaussian features with a range of spectral properties and durations. Among these, the most-often discussed are probably the noise transients (“glitches”) that haunt searches for compact-binary coalescences [35]. Yet, searches for continuous waves are most affected not by these short-lived

glitches, but rather by persistent narrow-band features (“lines”) [36]. Many of these spectral lines only become apparent after long periods of coherent observation, making their identification and eradication especially challenging. Noise-spectral lines could also be accompanied by sidebands which affects the sensitivity in a narrow-frequency region [37]. Furthermore, their distribution over the sensitive frequency band of the detectors is highly irregular and changes with time as the instruments evolve.

Fully coherent searches for continuous waves tend to have very high-frequency resolution (of order $\delta f \sim 10^{-7}$ – 10^{-8} Hz) scaling directly with the integration time ($\delta f \sim 1/T$). This fine resolution means that such analyses can fall victim to very narrow (and weak) noise lines. Furthermore, as mentioned at the end of Sec. II A 1, a pulsar signal will be spread over a band of width $\Delta f \approx f \times 10^{-4}$ Hz around its central GW frequency f . This means that attempts to find such a signal will be affected by noise over a range of frequencies, broad with respect to the typical resolution of the search. A persistent departure from Gaussianity in that frequency range (e.g., a wandering instrumental line that happens to intermittently cross the targeted band) will confound most searches, potentially yielding false positives (“outliers”). Naturally, the number of outliers due to unmodeled noise found by the pipelines will increase with the searched CW parameter space, as well as with observation time (which increases the frequency resolution and the sensitivity to unmodeled noise sources).

As is the case with the glitches affecting searches for compact binaries, lines and other non-Gaussian features would not be an issue for continuous-wave searches if there existed a robust way to model them and directly incorporate that knowledge into the statistical analyses (cf. Sec. II C below). However, the noise artifacts in the set that interests

us are, by definition, impossible to fully model from first principles: Any particular noise source that *is* well understood can usually be physically or digitally removed, so that they are no longer of concern [36,38–40]; the remaining artifacts are, therefore, those that are intrinsically unpredictable or so far not understood. Consequently, we are left to try to find ways to *empirically* estimate the true statistical background (i.e., the probability distribution of false positives) of a search in order to assign significances to potential detection candidates.

C. Searches

Searches targeted at known pulsars make use of the simple form of the expected signal Eq. (1) to match filter the data and determine the likelihood that a signal is present. There exist three standard approaches of this kind: the time-domain Bayesian [22,30,41,42] and \mathcal{F}/\mathcal{G} -statistic [23,43–45] methods, and the frequency-domain $5n$ -vector method [46–48]. Due to the technical details underlying each implementation, only the Bayesian time-domain method has been broadly applied to a large number of targets [16]. Although sky shifting is applicable to all three of these techniques, in the following sections we will only use the Bayesian and five-vector searches for concrete examples. For completeness, here we provide a brief overview of the basics of all three approaches.

1. Bayesian approach

Bayesian statistics provide a complete and straightforward framework for computing the probability that a given set of data contains a signal vs Gaussian noise, and for inferring the parameters that best describe the signal if present. The implementation utilized for searches by the LIGO Scientific Collaboration and Virgo Collaboration [22] takes advantage of the fact that the phase evolution $\phi(t)$ is known from electromagnetic observations to remove the high-frequency components of the signal early in the process—this dramatically simplifies the Bayesian inference step itself [30,42].

First, the data are digitally heterodyned [30,41], so that the signal they putatively contain becomes

$$h'(t) \equiv h(t)e^{-i\phi(t)} = \Lambda(t) + \Lambda^*(t)e^{-i2\phi(t)}, \quad (11)$$

with $*$ indicating complex conjugation, and

$$\Lambda(t) \equiv \frac{1}{4}F_+(t)h_0(1 + \cos^2 t) - \frac{i}{2}F_\times(t)h_0 \cos t. \quad (12)$$

A series of low-pass filters are then applied to remove the second term in Eq. (11), which enables the downsampling of the data by averaging over minute-long time bins. As a result, $\Lambda(t)$ is the only contribution from the original signal left in our binned data \mathbf{B} , which will now look like

$$B_{\text{expected}}(t_k) = \Lambda(t_k) + n(t_k), \quad (13)$$

where $n(t_k)$ is the heterodyned, filtered, and downsampled noise in bin k , which carries no information about the GW signal.

Equation (13) implies that the quantity $B(t_k) - \Lambda(t_k)$ should have the statistical properties of noise, and that Eq. (12) should be the template in our search. This knowledge can be used to compute the marginalized-likelihood ratio (Bayes factor) that the data contain a signal buried in noise (\mathcal{H}_S), vs just Gaussian noise (\mathcal{H}_N):

$$\mathcal{B}_N^S = \frac{P(\mathbf{B}|\mathcal{H}_S)}{P(\mathbf{B}|\mathcal{H}_N)}. \quad (14)$$

If the detector noise was indeed Gaussian, this single quantity would suffice to define a detection criterion: A value greater than unity would indicate the signal model is favored by that factor (in terms of betting odds), and vice versa. However, since actual noise cannot be guaranteed to be Gaussian (and, generally, will not be), the probability ratio of Eq. (14) does not inform us about the relative likelihoods of a signal vs *actual* (non-Gaussian) noise. To address this, one may attempt to capture instrumental artifacts by defining a construction similar to Eq. (14) but using signal coherence across detectors to distinguish spurious effects from actual astrophysical signals [16,49,50]. Nevertheless, it cannot be shown that any such construction will always capture all the features of real instrumental noise (in the language of formal logic, our hypothesis set is never complete). Therefore, we would benefit from a method to empirically test the efficacy of our Bayesian constructions at actually distinguishing signals from (non-Gaussian) detector noise.

2. Five-vector approach

The frequentist five-vector method [46] builds a detection statistic using the sidereal modulation given by the interferometer antenna response to the two CW polarizations, encoded by $F_{+/\times}$ in Eq. (1). Similar to the procedure outlined in Sec. II C 1, the first step is to remove all the possible phase modulations, apart from the sidereal ones caused by the antenna patterns. Depending on the type of search, this may be achieved through different techniques, including subheterodyning, nonuniform resampling, or a combination thereof [51–53]. After this step, the signal can be modeled via two sidereal responses $A_{+/\times}(t)$ analogous to $F_{+/\times}(t)$ but which do not depend on the polarization angle ψ (see [46] for more details). Concretely, it may be shown that $A_+(t) \equiv F_+(t; \psi = 0)$ and $A_\times(t) \equiv F_\times(t; \psi = \pi/4)$. Doing this, the signal assumes the complex-valued form:

$$h(t) = H_0(\eta)[H_+(\psi, \eta)A_+(t) + H_\times(\psi, \eta)A_\times(t)], \quad (15)$$

where η is related to the ratio of the two polarization amplitudes given in Eq. (1),

$$\eta = -\frac{2 \cos \iota}{1 + \cos^2 \iota} \quad (16)$$

and with $H_{+/\times}$ defined by

$$H_+ = \frac{\cos(2\psi) - i\eta \sin(2\psi)}{\sqrt{1 + \eta^2}},$$

$$H_\times = \frac{\sin(2\psi) - i\eta \cos(2\psi)}{\sqrt{1 + \eta^2}}.$$

Just as in the case of $\Lambda(t)$ in Eq. (12), the frequency components of a signal described by Eq. (15) are simply those corresponding to the sidereal modulations encoded in $A_{+/\times}(t)$. These frequency components ($f_{5\text{-vec}}^i$) are integer multiples of the sidereal rotation frequency of Earth, namely,

$$f_{5\text{-vec}}^i = f_{\text{gw}} + 2\pi k^i \omega_{\text{sid}}, \quad \vec{k} = [-2, -1, 0, 1, 2]. \quad (17)$$

Therefore, any signal like Eq. (15) may be described as a vector in the space spanned by the five δ -functions corresponding to the frequencies in Eq. (17).

To search for signals, the frequency-domain GW data can be projected onto the five-vector space, to obtain a set of projections \vec{X} . This resulting vector now lives in the same space as the sidereal templates, which can be represented as five-vectors $\vec{A}_{+/\times}$. We may thus obtain the matched filter between the data and the antenna patterns by taking a simple scalar product between \vec{X} and $\vec{A}_{+/\times}$. By maximizing this matched filter, one obtains an estimator for the GW polarization amplitudes:

$$\hat{H}_{+/\times} = \frac{\vec{X} \cdot \vec{A}_{+/\times}}{|\vec{A}_{+/\times}|^2}, \quad (18)$$

which can be in turn used to define a detection statistic

$$S_5 \equiv |\vec{A}_+|^4 |H_+|^2 + |\vec{A}_\times|^4 |H_\times|^2. \quad (19)$$

After carrying out the above procedure for templates corresponding to different parameters, detection candidates (i.e., values of the parameters for which might match a potential signal) are identified by their value of S_5 . In particular, the statistic is required to exceed a threshold corresponding to a preset false-alarm probability. To do this, one must know or measure the distribution of S_5 over noise. Traditionally this has been computed analytically by assuming purely Gaussian noise with known variance [46].

Alternatively, since real data are not Gaussian, one may try to approximate the background distribution by

computing S_5 over frequency bands far from the expected signal (off-frequency analysis). The frequency regions should be far enough from a possible CW signal such that only the noise contribution is present in the detection statistic, and close enough to the analyzed band to share its statistical properties. Given that the noise is strongly frequency dependent, finding this sweet spot is far from trivial (if at all possible) and one can never guarantee that the conditions required for an unbiased estimation of the background are being satisfied.

3. \mathcal{F} -statistic

The \mathcal{F} -statistic was first introduced in [23] for gravitational-wave searches from neutron stars, and was later extended for other astrophysical objects (e.g., [54,55]). In the case of Gaussian noise, the \mathcal{F} -statistic is defined as the natural logarithm of the maximum-likelihood ratio between the signal and noise hypotheses:

$$\mathcal{F} = \max_{\vec{\theta}} \left[\ln \frac{P(\mathbf{d}|\vec{\theta}, \mathcal{H}_S)}{P(\mathbf{d}|\mathcal{H}_N)} \right], \quad (20)$$

where \mathbf{d} are usually calibrated detector data, and the maximization is over the signal-template parameters, $\vec{\theta}$. It can be shown that the \mathcal{F} -statistic can be analytically maximized over the “extrinsic parameters” (h_0 , ψ , ι , and ϕ_0), thus reducing the dimensionality of the numerical computations to the so-called “intrinsic parameters” (α , δ , f , and \dot{f}). Since in a targeted search the intrinsic parameters are supposed to be perfectly known, a targeted search based on the \mathcal{F} -statistic would reduce to the computation of one value for \mathcal{F} , that is later compared to the expected noise-only distribution for Gaussian noise in order to assign the p -value.

The analysis proceeds by match filtering the data against four different templates, each of them corresponding to a particular combination of intrinsic phase evolution and sidereal modulation. The outcome of these four filters is the \mathcal{F} -statistic, for which, if the data are composed purely of Gaussian noise, it can be shown that the value $2\mathcal{F}$ follows a χ^2 distribution with 4 degrees of freedom [23]. Detection candidates (“outliers”) are selected according to their false-alarm probability, which can be computed analytically if one assumes Gaussian noise. However, false-positive outliers arise when the noise is not Gaussian, and thus not properly handled by Eq. (20). If this is the case and one cannot trust the background distribution of the statistic to be simple χ^2 , this distribution must be estimated empirically by producing sets of data known with certainty to contain no astrophysical signals.

III. METHOD

Having reviewed the basics behind targeted searches for continuous waves, including the difficulties inherent to

nonidealized instrumental noise, we here introduce sky shifting as a way to empirically assign detection significances. In Sec. III A we describe the basic ideas behind this simple procedure and explain how it can be easily applied to the Bayesian and five-vector searches. In Sec. III B, we heuristically explore the limits of applicability of this technique, concluding that sky shifting is a viable method for estimating the background distribution of detection statistics, as long as a few simple conditions are satisfied. This will be demonstrated in the following section (Sec. IV) with concrete examples.

A. Sky-shifting

Lacking a satisfactory way to model all noise artifacts and their effect on CW searches from first principles, we may instead attempt to empirically determine the distribution of the different search statistics in response to real detector noise and in the absence of astrophysical signals. As discussed in Sec. I, a naive attempt at blinding the data to astrophysical CWs using methods analogous to those used for compact-binary coalescences (CBCs) is doomed to fail. Therefore, we may instead look for a solution in specific properties exclusive to real gravitational signals, as opposed to noise.

One example of such a feature is the requirement of consistency between the phase evolution observed by EM astronomers, and the sky location of the source: While the two properties, as encoded in the signal itself, must necessarily agree in the case of a real GW, there is no special link between them in the case of noise artifacts. Furthermore, as explained in Sec. II A, the location of the source is independently imprinted in the morphology of the signal twice: In the amplitude modulation due to the antenna patterns, Eq. (1), and in the frequency modulation due to Rømer and other timing delays, Eq. (5). Since these three properties (frequency, amplitude modulation, and phase modulation) must all agree for an astrophysical signal, we may ask: How likely is it for an instrumental artifact to randomly satisfy this condition and thus mimic a real signal from a given source?

A priori, an instrumental artifact with frequency close to that expected from a given source is no more likely to also show the amplitude and phase modulations corresponding to the true location of the pulsar than those of any other arbitrary sky location. In other words, there is no reason for instrumental noise at the target frequency to “know” what the true sky location of the source is. By carrying out our analysis assuming incorrect sky locations, we may blind ourselves to astrophysical signals and empirically estimate the probability that instrumental artifacts in the narrow-frequency region corresponding to a given source also present the modulation matching its location in the sky.

The above idea may be rephrased in the language of function spaces. Continuous signals observed for a finite period of time can be represented as vectors in the space of

square-integrable functions (L^2) or, after discretization, the space of square-summable sequences (ℓ^2). We would like to estimate the overlap between the subspace of L^2 (or, rather, ℓ^2) occupied by noise features and the much narrower one spanned by the signal template, Eq. (1). We attempt to empirically achieve this by computing an inner product (defined by the detection statistic itself) between the data and signal templates (basis elements) corresponding to different sky locations. We expect this to work partly because templates for different sky locations will be morphologically very similar to the true template, while the same is not true for any arbitrary function. This also allows us to explore the statistical properties of the noise in the same region of frequency space occupied by the expected signal. (See also [56].)

An alternative to sky shifting would be to simply randomly shuffle the time series data used in the analysis; this would be easy to do for the time-domain data \mathbf{B} used in the Bayesian analysis, and is similar to previously implemented strategies (e.g., [57]). Such a shuffling would completely decohere any real signal in the data and, given the huge number of permutations of the shuffling, any random shuffling would be extremely unlikely to be correlated with any other. This approach would be the simplest way to proceed if the data were purely Gaussian and stationary across an observation run. However, for real data that contain contaminating instrumental lines and a time-varying noise level, it would also scramble these components. Therefore, any statistic produced with the scrambled data would not be a fair comparison to the unshuffled foreground. Shuffling would also be complicated to perform for a search method that begins with data that have already been transformed into the frequency domain. Sky-shifting, as described above, does not suffer from these limitations.

1. Implementation examples

Background distributions may be estimated via sky shifting in the context of any of the searches described in Sec. II C. This is true regardless of whether the search is carried out in the time or frequency domains, for one or several detectors, maximizing or marginalizing over nuisance parameters. This generality stems from the fact that sky shifting is largely insensitive to the specific details behind the computation, as long as the sky location and phase evolution are assumed to be known.

Let us first illustrate this by using the time-domain Bayesian search as a concrete example. As outlined in Sec. II C 1, this approach is split into two stages: (i) heterodyning of the data to put the signal in the shape of Eq. (12) and (ii) Bayesian inference to compute the relative likelihood of a signal being present, Eq. (14). It is important for our purposes that information about the location of the pulsar is only needed in the first step, making it straightforward to apply our suggested strategy. In particular, we may

intentionally heterodyne the data assuming an off-source (shifted) sky location, and then carry out the inference stage as usual, assuming the true (on-source) sky location. Rather than being indicative of a signal, a large Bayes factor obtained this way would necessarily reveal the presence of a noise artifact. This process may be repeated for different sky locations to obtain an estimate of how likely noise is to mimic a signal from this pulsar.

As another example, consider the frequency-domain five-vector approach of Sec. II C 2. In that case, we may also resample or reheterodyne the data assuming a shifted sky location during a preprocessing stage. This procedure is expected to spread the power of a possible GW signal over many different frequency bins, making it too weak to be detectable and thus blinding the analysis to it. Next, we compute the S_5 statistic by using the five-vector sidereal function $A_{+/\times}(t)$ computed for the on-source sky position. We can then repeat these steps for many different sky locations to obtain a collection of background values for the S_5 statistic. This yields a noise-only distribution for the S_5 statistic that quantifies the probability for a noise disturbance to mimic the sidereal antenna patterns corresponding to the true sky location.

B. Blinding and draw independence

In order for the method above to work, we need to make sure that the different sky locations used are actually distinct, so that there is no leakage of a possible GW signal in the draws used for building the noise-only statistic distribution. In order for the method above to work, we need to make sure that (i) the different sky locations used are actually distinct, so that the results can be treated as distinct draws from the probability distribution we are trying to estimate, and (ii) sky shifting really does blind the data to foreground signals. The first requirement is easy to satisfy and translates into the need for picking sky locations with angular separations greater than the worst (largest) sky bin resolvable by the search, as explained in Sec. II A 2. As we show below, the second requirement can also be satisfied by picking off-source locations far enough away from the true position of the source.

1. Signal contribution to sky-shifted statistic

Let us first examine the conditions under which sky shifting effectively removes contributions from real continuous waves. For simplicity, consider a signal of fixed frequency (f) originating from some known location ($\hat{\Omega}_{\text{on}}$). Now imagine heterodyning the data containing it by using a mismatched timing correction corresponding to some shifted sky location ($\hat{\Omega}_{\text{off}}$), as proposed in Sec. III A. In full analogy to Eq. (13), we would then obtain binned data like

$$B(t) = \Lambda(t)e^{2\pi if\Delta\tau(t)} + n(t) \equiv \Lambda'(t) + n(t), \quad (21)$$

where we no longer assume the instrumental noise $n(t)$ is normally distributed, and where

$$\Delta\tau(t; \Delta\hat{\Omega}) \equiv \tau(t; \hat{\Omega}_{\text{on}}) - \tau(t; \hat{\Omega}_{\text{off}}) \quad (22)$$

represents the timing-correction mismatch between the two sky locations with angular separation $\Delta\hat{\Omega} = \hat{\Omega}_{\text{off}} - \hat{\Omega}_{\text{on}}$ [cf. Eq. (5)]. As a proxy for a generic search statistic, consider the evaluation of a simple inner product between the data and the expected template:

$$\langle B(t)|\Lambda(t) \rangle = \langle n(t)|\Lambda(t) \rangle + \langle \Lambda'(t)|\Lambda(t) \rangle. \quad (23)$$

Our goal is to estimate the distribution of the overlap between the noise and the template $\langle n(t)|\Lambda(t) \rangle$ by studying our proxy statistic $\langle B(t)|\Lambda(t) \rangle$. Consequently, we would like the contribution of the true signal $\langle \Lambda'(t)|\Lambda(t) \rangle$ to Eq. (23) to be sufficiently small to be effectively undetectable.

Explicitly, the contribution of the signal to the inner product of Eq. (23) can be written in terms of a time integral over the observation time T ,

$$\langle \Lambda'(t)|\Lambda(t) \rangle = \left| \int_0^T \Lambda^2(t) e^{2\pi if\Delta\tau(t; \Delta\hat{\Omega})} dt \right|, \quad (24)$$

where $\Lambda^2(t) \equiv \Lambda^*(t)\Lambda(t)$. The first key feature of this result is that a signal with greater signal-to-noise ratio (SNR) will tend to contaminate the sky-shifted statistic more strongly. This is not at all surprising: A strong signal can be detected even if there is a small error in its assumed sky location, because enough coherent power can remain even after the timing correction spreads it over several frequency bins. In fact, Eq. (24) is bounded from above by the squared norm of the signal template,

$$\langle \Lambda'(t)|\Lambda(t) \rangle \leq \langle \Lambda(t)|\Lambda(t) \rangle = \left| \int_0^T \Lambda^2(t) dt \right|, \quad (25)$$

which, for a flat power spectrum, is directly proportional to the square of the SNR.

The second relevant feature of Eq. (24) is that the dependence of $\langle \Lambda'(t)|\Lambda(t) \rangle$ on sky location will be determined solely by the angular structure of $\Delta\tau$, and how well that can be resolved given f and T (see Sec. II A 2). The inequality of Eq. (25) is, of course, saturated if and only if the shifted location is such that $\Delta\tau(t; \Delta\hat{\Omega}) = 0$ at all times: This takes place, for instance, if the “shifted” location is really just the original location of the source ($\Delta\hat{\Omega} = 0$). On the other hand, for most other values of $\Delta\hat{\Omega}$ and for the range of f ’s we are interested in, the exponential term in Eq. (24) is highly oscillatory—this means that we should expect $\langle \Lambda'(t)|\Lambda(t) \rangle$ to quickly vanish as we move away from the true location of the source. This is consistent with the sky-bin definition given in Sec. II A 2, from which it is possible to see that the sky-bin size decreases with

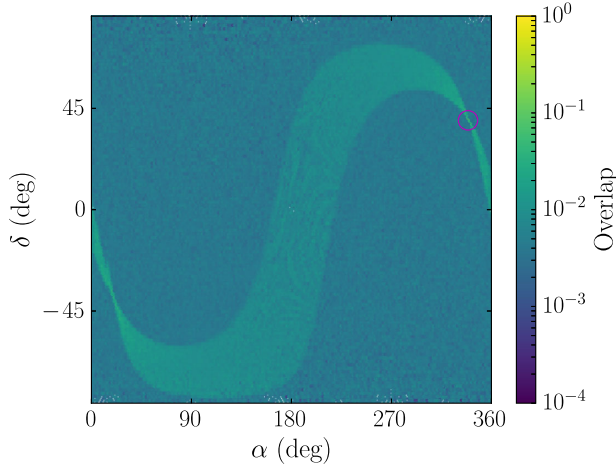


FIG. 2. Value of the reheterodyned correlation overlap of Eq. (26) over the entire sky, as a function of right ascension (α) and declination (δ). The data contain a simulated signal with SNR = 70 located at $\alpha = 22^{\text{h}}35^{\text{m}}40.73^{\text{s}}$, $\delta = 39^{\circ}40'44.76''$ (magenta circle).

increasing frequency. In other words Eq. (24) is representative of the templates density lattice for a GW search [58–60].

2. Angular pattern and magnitude

The detailed angular structure of our proxy sky-shifted statistic is represented in Fig. 2. To produce this plot, we began with a set of binned data, Eq. (13), containing Gaussian noise and a very strong (SNR = 70) simulated signal from an arbitrary sky location on the ecliptic plane (indicated by a magenta circle). We then reheterodyned these data assuming different (sky-shifted) locations covering the whole sky, and for each instantiation computed the overlap (normalized cross-correlation),

$$\text{Overlap} = \frac{\langle B'(t)|B(t) \rangle}{\langle B(t)|B(t) \rangle} \approx \frac{\langle \Lambda'(t)|\Lambda(t) \rangle}{\langle \Lambda(t)|\Lambda(t) \rangle}, \quad (26)$$

between the sky-shifted data $B'(t)$ and the on-source data $B(t)$. This quantity (shown in color in Fig. 2) represents the normalized contribution of the injected signal to the sky-shifted statistic for different sky locations, as desired. This is because $\langle n'(t)|n(t) \rangle \approx 0$ for Gaussian noise, yielding the approximate equality in Eq. (26).

As expected, the contribution of the signal falls off steeply as we move away from the source location: While the normalized overlap of Fig. 2 equals unity if $\Delta\hat{\Omega} = 0$ (center of the magenta circle), it is orders of magnitude smaller for all other choices of $\hat{\Omega}_{\text{off}}$. The rest of the structure in this plot reflects the symmetries of the timing correction Eq. (5), which are themselves dominated by the Rømer term in Eq. (6): Locations across lines of fixed ecliptic latitude remain somewhat correlated to the on-source

location, and the whole pattern is symmetric under reflections through the ecliptic (see Sec. II A 2 for more details). This important observation means that, for any given on-source location, we will want to sample our sky-shifted points from only one of the ecliptic hemispheres. This may be achieved by laying out a grid in the sky, or by picking sky locations randomly. In either case, one must ensure uniqueness of the chosen point by enforcing a minimal separation set by the overlap function in Fig. 3.

For a sufficiently loud signal, sky bins neighboring the source will yield contaminated sky-shifted data (i.e., data that still contain measurable coherent power due to the on-source signal). Unlike the angular dependence, the overall magnitude of this contamination will be determined by the SNR of the signal and, as such, will depend on the integration time and intrinsic amplitude. For the same example as in Fig. 2, Fig. 3 shows the rate at which the overlap with the on-source location decreases as one moves at constant ecliptic latitude away from the source and for different integration times.

Because latitude is held constant in this plot, Fig. 3 represents the slowest-possible decrease in the contamination by this source (cf. Fig. 2). Furthermore, this example was chosen to have very high SNR and to lie on the ecliptic plane, where the sky resolution is poorest (cf. Sec. II A 2)—all of which makes this close to a worst-case scenario. Despite this, the overlap vanishes quite quickly, plateauing far away from the source at a value of the order 0.01. For realistic (low SNR) CW signals a 1% overlap is small enough to make the signal term in Eq. (23) negligible with respect to the noise. Hence for one month of data integration or more, setting sky-shifting separation to 1 deg is enough to remove any measurable correlation between sky bins in any realistic situation. Examples of this are given in Sec. IV below.

3. Contaminated backgrounds

An analysis that draws part of the background from a measurably contaminated region may underestimate the significance of the true signal, but never overestimate it. This is because a contaminated background will show artificial tails toward higher values of the detection statistic, due to coherent power left over after sky shifting in some of the “noise-only” instantiations. Thus, in a sense, such an analysis would, at worst, be conservative. Yet, as we will show in Sec. IV, a signal that is sufficiently loud to cause such contamination over a non-negligible region of the sky will itself yield an on-source detection statistic that is significantly higher than any of the contaminated-background tails. Therefore, the significance (e.g., p -value) assigned to such a signal will be the same with or without the tails.

In any case, sky bins in the immediate vicinity of the source may always be removed from the background estimation to prevent contamination. However, the excision

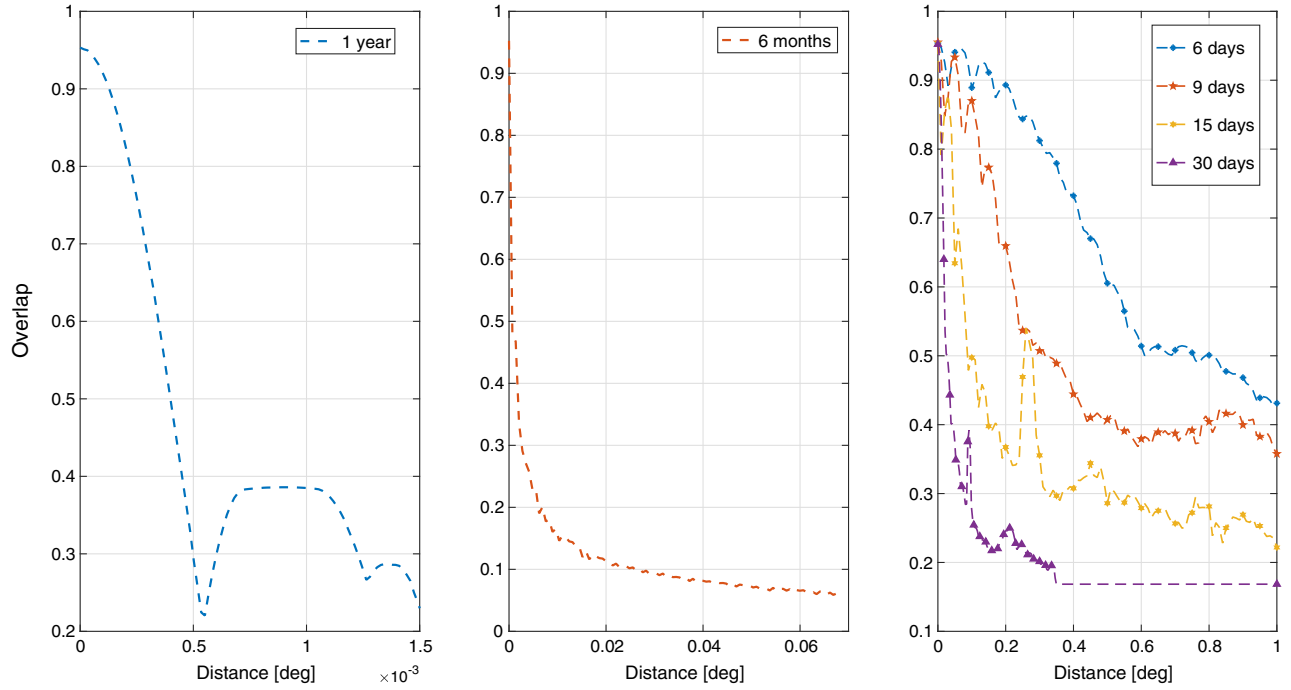


FIG. 3. Left panel: overlap Eq. (26) computed for a software injection with the same parameters as Fig. 2 for one year of integration time and sky locations close to the source; the overlap drops very rapidly over an angular distance of $\sim 5 \times 10^{-4}$ deg from the source, keeping ecliptic latitude fixed. Central panel: same as in the first panel but for six months of integration; the overlap drops over a greater distance than for one year of integration. Right panel: overlap functions for various integration times, as indicated by the legend; shorter integration times induce correlations over greater angular scales. The on-source overlap (null distance) is not identically 1 due to the presence of noise.

of a large area of the sky will have the detrimental consequence of effectively reducing the number of sky bins available for background estimation. Furthermore, such procedure is only justified if we (implicitly) assume that the on-source data do contain an astrophysical signal. In a way, this is analogous to how a very loud CBC signal may pollute the time-slid background in searches for transient gravitational waves (e.g., see caption to Fig. 3 in [61], or Fig. 7 in [62]). In that case, the standard procedure has been to first compute significances with the “polluted” background to determine whether the zero-lag detection candidate is a real signal, and only if that is the case remove it from the background.² The same can be done here if necessary.

In summary, we conclude that sky shifting, as described in Sec. III A, is a viable method for estimating the background distribution of detection statistics in targeted searches for continuous waves, as long as the shifted sky locations are chosen such that (i) they are distributed over only half the sky; (ii) the angular distance between them is no shorter than the sky resolution of the search

(cf. Sec. II A 2). This will guarantee that the different draws from the background distribution (obtained from different shifted sky locations) are distinct and uncontaminated by a true signal, if present.

IV. ANALYSIS

We study the efficacy of sky shifting (Sec. III A) as a viable method to empirically estimate the background distribution of detection statistics in targeted searches for continuous gravitational waves from known pulsars. We do this in the context of both the Bayesian (Sec. II C 1) and frequentist five-vector (Sec. II C 2) analyses, to demonstrate the generality of the approach. We discuss specific case studies in Sec. IV A, and systematically compare to different methods by computing false-dismissal and false-alarm rates in Sec. IV B.

The following results make use of both simulated and actual noise from interferometric detectors. In all cases, we begin with a set of data representing the (synthetic or actual) output of a detector after applying the preprocessing required to target some chosen pulsar (e.g., filtering and downsampling)—these are the on-source data. We then proceed as described in Sec. III A to generate multiple new sets of sky-shifted data, and then evaluate the distribution of the detection statistic over all such instantiations (excluding the original, on-source data). We can then compare the

²These two kinds of backgrounds are known colloquially as with and without “little dogs,” since this distinction first arose during the analysis of an injection in the direction of *Canis Major* [63].

TABLE I. Parameters for the case-study signals (Sec. IV A).

	f_{GW} (Hz)	α	δ	Figures	Data	Comment
J0534 + 2200	59.33	5 ^h 34 ^m 31.97 ^s	22°00′52.07″	5, 6	Gauss. design (H,L,V)	Assumed $\iota = 61.3^\circ$, $\psi = 124.0^\circ$
J1932 + 17	47.81	19 ^h 32 ^m 07.17 ^s	17°56′18.70″	7	Real O1 (H,L)	Published in [16]
P03	108.86	11 ^h 53 ^m 29.42 ^s	−33°26′11.77″	8	Real O1 (H,L)	Hardware injection [64]

value of the on-source statistic to the sky-shifted background, as we would in a real analysis.

A. Case studies

Here we provide several concrete examples of sky shifting at work in the presence of pure noise, realistic signals, and strong signals, as summarized in Table I. Background distributions are estimated from 10^4 shifted locations in the hemisphere of the source. The simulations of Gaussian noise (Sec. IV A 1) were carried out assuming an observation time of six months and power spectral densities (PSDs) corresponding to the aLIGO and Virgo design sensitivities. With the exception of Fig. 4, the simulated data for LIGO Hanford (“H”), LIGO Livingston (“L”), and Virgo (“V”) detectors were then analyzed coherently with the Bayesian method of Sec. II C 1, to obtain the signal vs noise Bayes factors of Eq. (14) as our detection statistic.

The examples with real instrumental noise correspond to LIGO’s first observation run (O1). The data streams start on September 11, 2015 at 01:25:03 UTC for Hanford and 18:29:03 UTC for Livingston, and finish on January 19, 2016 at 17:07:59 UTC at both sites. The first example consists of data prepared for the pulsar PSR J1932 + 17, for which search results were presented in [16]. All analysis settings are the same as in [16], except for a log-uniform prior in the signal amplitude (same as in [65]). The second example is for a hardware injection presented in [64]. Both these examples are offered merely to demonstrate the performance of sky shifting under realistic circumstances—we present no new observational results.

1. Pure-Gaussian noise

We first demonstrate that sky shifting works as expected in pure-Gaussian noise and in the absence of signal. In this case, the on-source data are just a set of samples from a Gaussian distribution with zero mean and standard deviation given by the value of the detector PSD at the GW frequency expected from the targeted pulsar. The sky-shifting process should correspondingly produce multiple instantiations of independent Gaussian noise, a fact that should be reflected in the resulting background distribution of the detection statistic. This distribution is shown in Fig. 4 for an example using the five-vector statistic of Eq. (19). When computed over Gaussian noise, it can be shown that this statistic must follow a Γ distribution with 2 degrees of freedom [46]. Figure 4 shows that this is the case, in agreement with our

expectation that sky shifting should produce independent draws from the background distribution.

2. Injections in Gaussian noise

Ideally, the background distribution should be unaffected by the presence of a signal: While the value of the on-source statistic should rise to reveal it, the sky-shifted values should be insensitive to it. We demonstrate that this is the case by injecting signals of different amplitudes in Gaussian noise. We simulate a signal from the Crab Pulsar (PSR J0534 + 2200) as seen by three advanced detectors (H, L, V) at design sensitivity over six months, and recover it using the Bayesian method of Sec. II C 1.

We first choose a realistic signal amplitude of $h_0 = 10^{-26}$, which is weak enough to be consistent with the latest upper limits for this source [20], but strong enough to yield a non-negligible network SNR = 5 for the chosen PSDs and observation time. The data containing the injected signal are then reheterodyned for 10^4 shifted sky locations to yield the survival function in Fig. 5, i.e., the

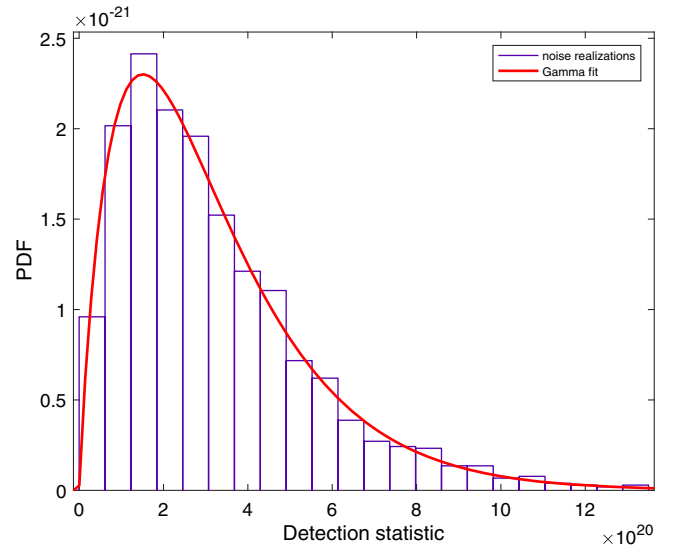


FIG. 4. Histogram of the noise-only distribution obtained using the sky-shifting method in case of Gaussian noise (no signal). Red line: best fit given by a Γ distribution with $a = 1.95$. In this test case the hypothetical source was assumed at a frequency $f \approx 108.85$ Hz with a spin-down $\dot{f} \approx 10^{-17}$ Hz/s. The on-source position was chosen on the ecliptic plane ($\alpha = 22^{\text{h}}35^{\text{m}}40.73^{\text{s}}$, $\delta = 39^\circ40'44.76''$) as in Fig. 2. The search assumed a single detector at the LIGO Hanford site.

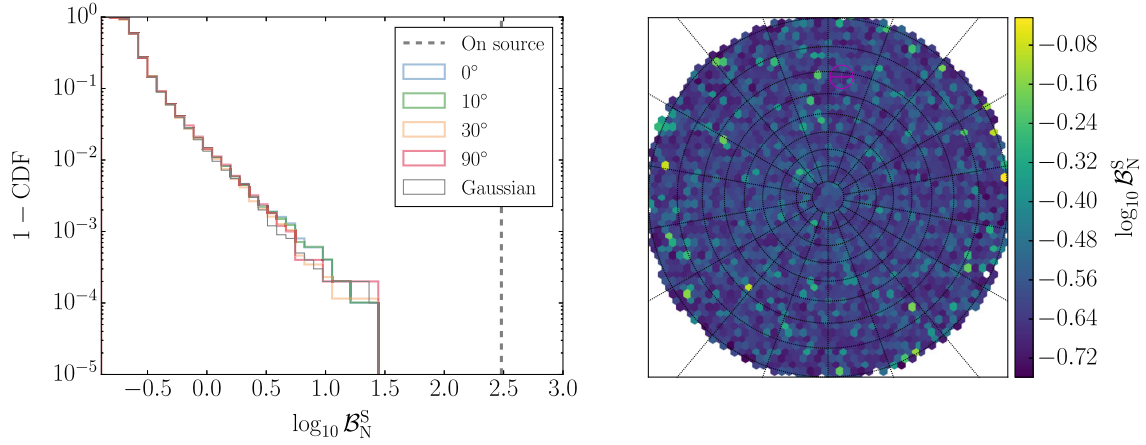


FIG. 5. Realistic Crab signal in Gaussian noise. We simulate a six-month-long signal with $h_0 = 10^{-26}$ (network SNR 5) and parameters consistent with the Crab Pulsar (PSR J0534 + 2200), inject it in Gaussian noise for aLIGO and Virgo design PSDs, and recover it using the Bayesian analysis of Sec. II C 1 (see Table I). The left panel shows the survival function (1-CDF) of $\log_{10} \mathcal{B}_N^S$, Eq. (14), for the sky-shifted background produced from the injected data for different excision areas around the source (different colors, blue and green overlap almost perfectly), as well as from pure-Gaussian noise (gray, thin histogram); the on-source statistic for the injection is $\log_{10} \mathcal{B}_N^S = 2.5$ (thick dashed line), higher than any of the 10^4 sky-shifted instantiations. The right panel shows the distribution of the sky-shift statistic over the sky in a north-polar stereographic projection, with the Crab's location marked by the magenta crosshairs; the color of each hexagon gives the average of $\log_{10} \mathcal{B}_N^S$ over several sky bins.

complement of the cumulative density function (CDF), 1-CDF. Each colored trace in this figure represents the distribution of sky-shift background computed from the northern celestial hemisphere, *excluding* any points closer to the source than the indicated angular distance, i.e., excluding points with $|\alpha - \alpha_\star| < \Delta$ and $|\delta - \delta_\star| < \Delta$ with a “ \star ” indicating the true location of the Crab and Δ one of the values given in the legend of Fig. 5: 0° (blue), 10° (green), 30° (yellow), or 90° (red). In particular, the blue curve corresponds to background from points sampled over the whole hemisphere, while the red curve corresponds to points sampled over the half-hemisphere not containing the source.

In this case, the choice of sky region does not have a strong effect on the background: We may take advantage of the whole hemisphere, getting quite close to the source (as allowed by the frequency resolution of this search). In fact, note that the blue and green curves in Fig. 5 are essentially identical. For reference, the distribution of the sky-shift statistic over the whole northern sky (0° curve on the left) is represented on the right panel of Fig. 5 via a stereographic map, with the true location of the source indicated by magenta crosshairs.

As expected, the background produced via sky shifting is practically indistinguishable from results in pure-Gaussian noise (gray, thin histogram). Indeed, these two samples yield a Kolmogorov-Smirnov (KS) p -value of 0.77, favoring the hypothesis that they were both drawn from a common distribution. This agreement is despite the fact that the on-source statistic (dark gray dashed line) takes a significantly increased value, revealing the presence of the injection. Completely ignoring the intrinsic probabilistic

meaning of \mathcal{B}_N^S , a background like the blue curve in the left panel of Fig. 5 would allow us to place a p -value of at most 10^{-4} on the null hypothesis that the on-source data are noise.

As anticipated in Sec. III B, there is a limit to how loud the injection can be without noticeably contaminating nearby sky bins and, therefore, biasing the background distribution obtained through sky shifting. However, this threshold is quite high: For the same detector configuration as above, we find that the injection must reach $h_0 \sim \mathcal{O}(10^{-24})$, or a network SNR ~ 700 at design sensitivity, before sky shifting is unable to effectively remove it. We show an example of this in Fig. 6 for a signal from the Crab Pulsar with $h_0 = 1.4 \times 10^{-24}$, which roughly corresponds to the spin-down limit for this source [16].³ This time, as seen from the panel on the left, the full-hemisphere sky-shifting distribution (blue curve) is visibly inconsistent with a pure-Gaussian background (gray, thin curve), and a KS p -value of 10^{-7} strongly disfavors a shared distribution between the two sample sets. From the right panel, it is clear that the culprits are noticeably contaminated sky locations in the neighborhood of the source (magenta crosshairs). These polluted sky bins are arranged in the same pattern predicted in Fig. 2, although under a different guise due to the logarithmic color scale.

³For an isolated pulsar (no accretion), the spin-down limit is the maximum power that could possibly be emitted in gravitational waves given the observed decay in the pulsar's angular momentum.

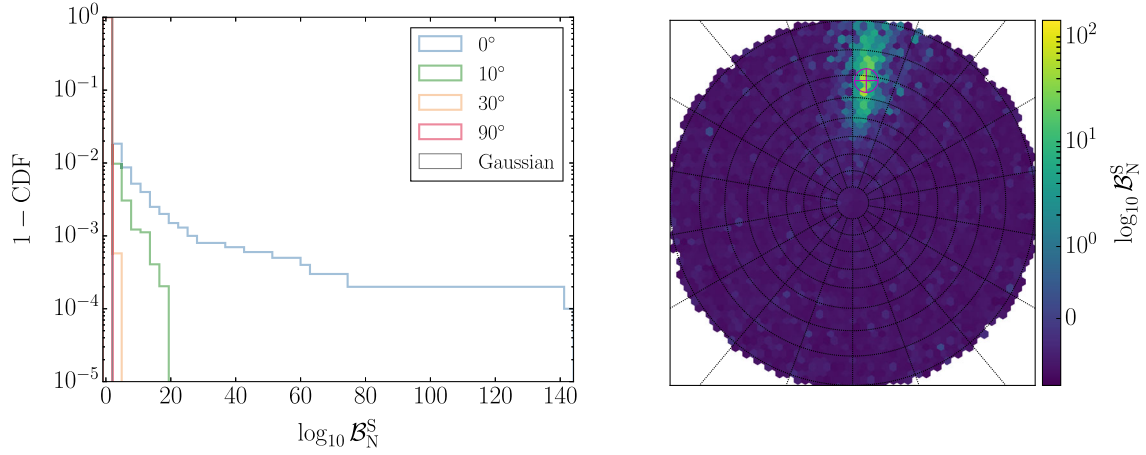


FIG. 6. Very loud Crab signal in Gaussian noise. We simulate a six-month-long signal with $h_0 = 1.4 \times 10^{-24}$ (network SNR 700) and parameters consistent with the Crab pulsar (PSR J0534 + 2200), inject it in Gaussian noise for aLIGO and Virgo design PSDs, and recover it using the Bayesian analysis of Sec. II C 1 (see Table I). The left panel shows the survival function (1-CDF) of $\log_{10} \mathcal{B}_N^S$, Eq. (14), for the sky-shifted background produced from the injected data for different excision areas around the source (different colors), as well as from pure-Gaussian noise (gray, thin histogram); the on-source statistic for the injection is $\log_{10} \mathcal{B}_N^S = 9 \times 10^4$, which is vastly higher than any of the 10^4 sky-shift instantiations (off the scale). The right panel shows the distribution of the sky-shifted statistic over the sky in a north-polar stereographic projection, and with the Crab's location marked by the magenta crosshairs; the color of each hexagon gives the average of $\log_{10} \mathcal{B}_N^S$ over several sky bins, in semilog scale linearly interpolated between $(-1, 1)$.

As discussed in Sec. III B, background contamination can at worst cause us to *underestimate*, never overestimate, the significance of a detection. Yet, the background is not underestimated in the example of Fig. 6 because the signal is too loud ($\log_{10} \mathcal{B}_N^S = 9 \times 10^4$, off the scale of the histogram in Fig. 6). This is a general feature: In Gaussian noise, if a signal is loud enough to contaminate a large region of the sky, it will also be louder than the loudest background produced from it.

3. Real noise

The above behavior is replicated in the presence of actual noise from LIGO and Virgo, with the difference that the background naturally shows tails due to the non-Gaussianities in the data. An example of this is shown in Fig. 7, which was produced using actual data from aLIGO's first observation run prepared for the pulsar PSR J1932 + 17 and with both detectors analyzed coherently using the Bayesian method of Sec. II C 1. As before, the left panel shows the sky-shifted background distributions for

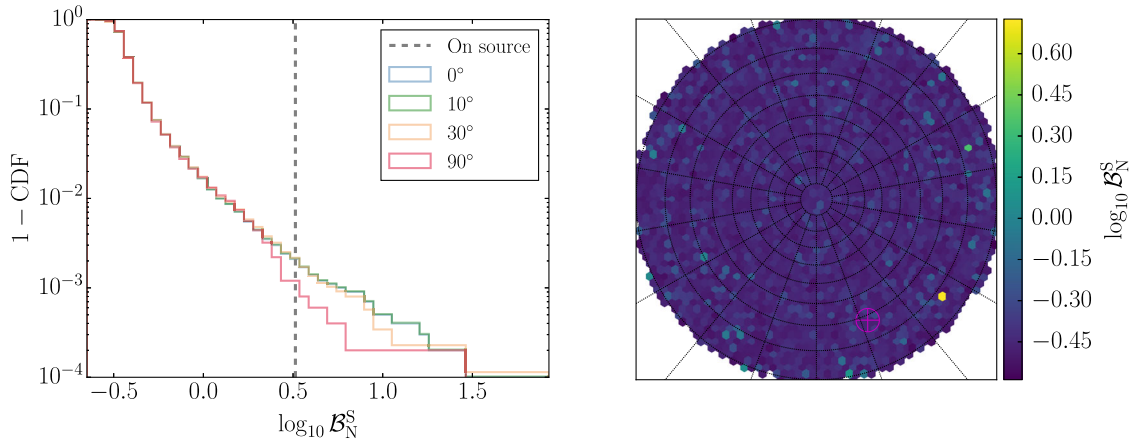


FIG. 7. Real O1 noise for J1932 + 17. Off-sourced background produced from real O1 LIGO data prepared for PSR J1932 + 17, analyzed coherently with the Bayesian method of Sec. II C 1 (see Table I). The left panel shows the survival function (1-CDF) of $\log_{10} \mathcal{B}_N^S$, Eq. (14), for the sky-shifted background for different excision areas around the source (different colors); the on-source statistic for this pulsar is $\log_{10} \mathcal{B}_N^S = 0.5$ (vertical dashed line), which differs from the result in [16] only due to a log-uniform prior on the signal amplitude. The right panel shows the distribution of the sky-shifted statistic over the sky in a north-polar stereographic projection, with the true location marked by the magenta crosshairs; the color of each hexagon gives the local average of $\log_{10} \mathcal{B}_N^S$ over several sky bins.

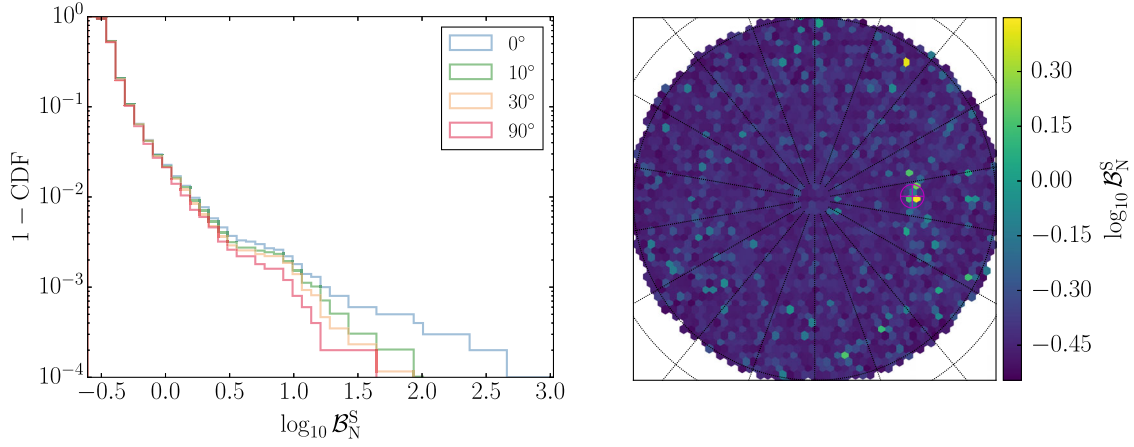


FIG. 8. Real O1 noise for a loud hardware injection (P03). Off-sourced background produced from real O1 LIGO data prepared for hardware injection P03 [64] analyzed coherently with the Bayesian method of Sec. II C 1 (see Table I). The left panel shows the survival function ($1 - \text{CDF}$) of $\log_{10} \mathcal{B}_N^S$, Eq. (14), for the sky-shifted background for different excision areas around the source (different colors); the on-source statistic for this pulsar is $\log_{10} \mathcal{B}_N^S = 504$ (off scale). The right panel shows the distribution of the sky-shift statistic over the sky in a south-polar stereographic projection, with the true location marked by the magenta crosshairs; the color of each hexagon gives the local average of $\log_{10} \mathcal{B}_N^S$ over several sky bins.

different excision areas around the source (different colors). Note that the excision process does not have any significant impact on the distribution, which is what one would expect in the absence of a very loud signal at the on-source location. The presence of artifacts in the data becomes apparent in the slower drop of the survival function with respect to, e.g., Fig. 5. The on-source value of the signal vs noise Bayes factor for this source was published in [16], and is marked here by a vertical dashed line—clearly, there is no evidence for a signal in the data.

To study the effectiveness of sky shifting in detecting a signal in real noise, we analyze data for the hardware injection referred to as “P03” in [64]. Hardware injections are produced by physically actuating on the test masses to mimic the effect of a true gravitational wave, providing a valuable end-to-end test of the instrumental calibration and analysis pipelines. In the case of P03, the signal was injected at 108.86 Hz with an amplitude of $h_0 = 8.2 \times 10^{-25}$ (network SNR = 50). As shown on the left panel of Fig. 8, this signal seems to be sufficiently strong to slightly contaminate the sky bins in its immediate vicinity, but this pollution is easily removed via a narrow excision (compare the blue trace to the rest in Fig. 8). In any case, the value of the on-source Bayes factor for this signal is $\log_{10} \mathcal{B}_N^S = 504$, which is significantly louder than the loudest background. Given that 10^4 sky-shifted noise instantiations were used to estimate the background, this implies that Fig. 8 would allow us to claim a detection of P03 with $p \leq 10^{-4}$ (ignoring the intrinsic probabilistic meaning of \mathcal{B}_N^S).

B. Comparison to other methods

In order to determine whether sky shifting offers an improvement over other strategies, we must go beyond

specific examples and study false-alarm and false-dismissal rates. That is, respectively, how likely is sky shifting to conclude that a noise artifact is a signal (false alarm), versus how likely is it to conclude that a signal is a noise artifact (false dismissal), as a function of the confidence level. We estimate those rates from a large number of injections in simulated and actual noise, and use them to directly compare with the standard background-estimation method for the five-vector search (Sec. II C 2). In our simulations, we find that sky shifting can outperform the usual methods in real LIGO data.

1. False-dismissal rate

First, in order to study the false-dismissal rate, for a selection of SNRs, we simulate 250 signals over the sky, with extrinsic parameters (ψ, η, ϕ_0) picked randomly over their allowed ranges. We then inject these in both idealized (Gaussian) and actual O1 noise for the LIGO Hanford and Livingston detectors (four months observation time). In the case of Gaussian noise, we pick the injection frequency and spin-down to be equal to those of P03, a choice that will only affect the specific size of the sky patches that required for sky shifting per Eq. (9). When using real detector data, we set the frequency of the injections to be 54.5 Hz with no spin-down. This is because O1 data are known to be polluted by noise artifacts in this frequency band, especially in the Livingston detector [66], making this a good frequency region to test the performance of sky shifting under realistic circumstances.

In each case, we use the method of Sec. II C 2 to obtain the on-source value of the detection statistic, as well as 2×10^4 sky-shifted background values with a minimal separation of 0.01 deg. We then compute the number of

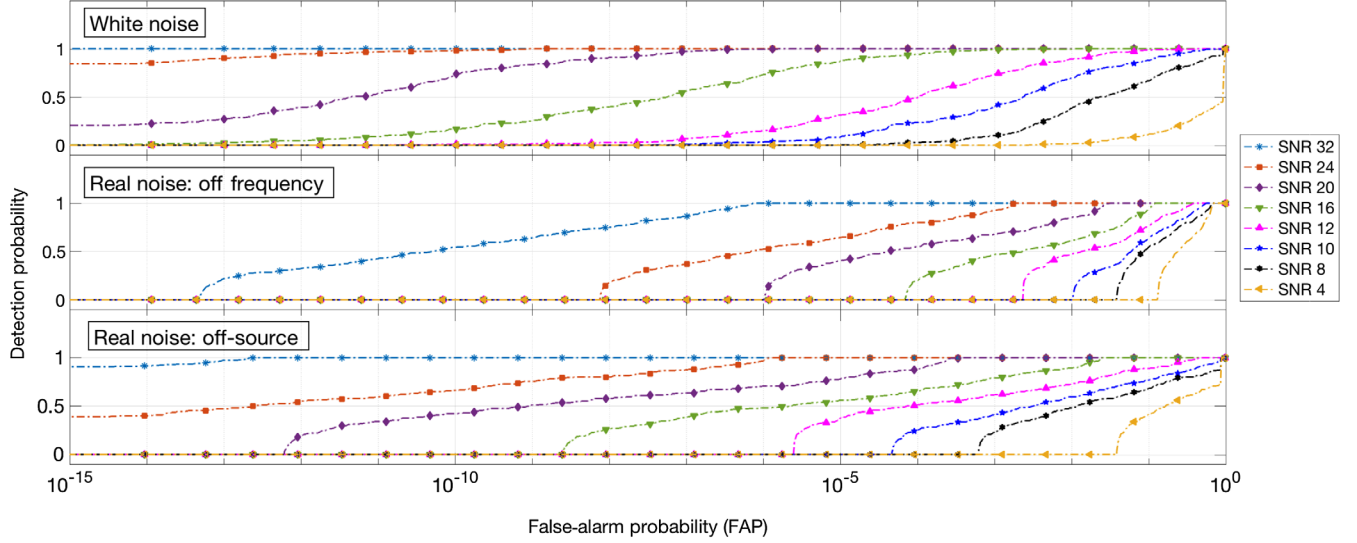


FIG. 9. Top panel: detection probability vs FAP threshold in Gaussian noise. Middle panel: same but done in non-Gaussian noise using the off-frequency method. Bottom: same again but this time using the sky-shift method.

detected signals as a function of the false-alarm probability, i.e., the number of injections recovered with a detection statistic that is higher than or equal to the value corresponding to a certain p -value, as established from the empirically estimated background. For comparison, we repeat the above procedure but with a background generated via the standard off-frequency method (mentioned in Sec. II C 2 and described in detail in [66]), instead of sky shifting. In both cases, the tails of the background distributions are extrapolated for very large values of the detection statistic using an exponential-decay fit.

The results of this study are summarized in the receiver operating characteristic (ROC) curves of Fig. 9. Curves in that figure represent the normalized detection rate (detection probability) as a function of the p -value (FAP) for different choices of injected SNR (different colors and traces) and different methods used to estimate the false-alarm probability (FAP). The top panel shows the results from applying the sky-shifting method to the case of Gaussian noise, the middle panel from applying the off-frequency method to real detector data, and the bottom panel from applying the sky-shifting to the same real data as in the middle plot. As expected, optimal ROCs are obtained in the case of Gaussian noise, for which both methods are equivalent; on real detector data, however, sky shifting shows better ROCs.

Non-Gaussian artifacts in real noise produce tails in the background distribution, hurting our chances to detect actual signals. This can be seen by comparing Fig. 5–8. This shows in Fig. 9 which indicates that, with those settings, we are only 50% likely to detect a SNR = 16 signal (green curve, down-pointing triangles) with FAP = 10^{-6} ($\sim 5\sigma$) in real O1 noise, but we are 75% likely to detect it with the same confidence in Gaussian noise.

Nevertheless, Fig. 9 also shows that sky shifting can outperform traditional methods in the presence of actual instrumental noise, as can be seen by comparing the center and bottom panels. For instance, the off-frequency method (middle panel) has essentially 0% chance of detecting a SNR = 16 signal at FAP = 10^{-6} , which is dramatically less than the 75% chance of detecting it via sky shifting (bottom panel). In fact, for these settings, sky shifting is consistently superior at all SNRs.

2. False-alarm rate

Besides assigning high significance to real signals, sky shifting should also be able to robustly reject outliers arising from non-Gaussianities in the data. In other words, it should have a low false-alarm rate at any given level of confidence, rejecting artifacts with high probability. The study of noise features is a very general problem due to the wide morphological range of non-Gaussianities that can be found in the data, which is the same reason why modeling noise likelihoods from first principles is impossible in the first place. This makes the benchmarking of noise-rejection probability a challenging problem.

To address this, we use as proxy simulated monochromatic noise lines at frequencies close to the targeted pulsar frequency. The putative source for which we were looking was again a pulsar with the rotational parameters of P03 and the sky position $\alpha = 22^{\text{h}}35^{\text{m}}40.73^{\text{s}}$, $\delta = 39^{\circ}40'44.76''$. In particular, we produce 300 datasets with noise lines added to four months of Gaussian noise with varying SNRs. Each noise line is injected with a frequency within 0.01 Hz of the targeted frequency, ensuring that the Rømer correction will cause it to contaminate the on-source analysis. This is because, for a putative source at

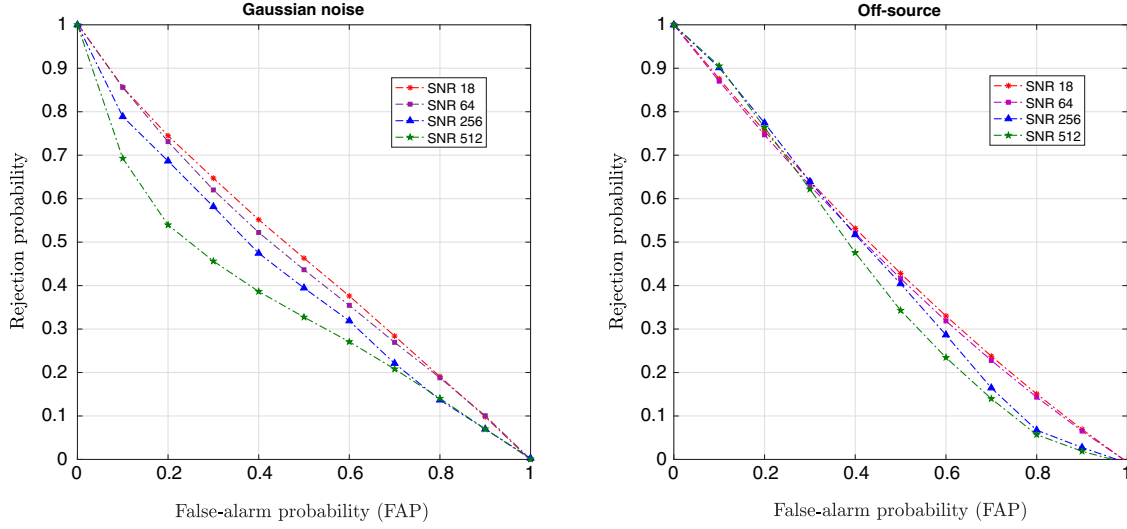


FIG. 10. Rejection probability of an outlier arising from persistent monochromatic noise lines in raw data. Significance for selection estimated using a Gaussian model for the noise (left) and the sky-shift method (right). The ordinate shows the percentage of noise lines that generate an outlier with a significance greater than or equal to the threshold corresponding to the FAP in the abscissa.

$f = 108.85$ Hz (see Table I), the Rømer frequency shift will be at most [51]

$$\Delta f = f \times 10^{-4} \approx 0.01 \text{ Hz}. \quad (27)$$

We pick the specific frequency and phase of the noise lines from a uniform distribution. After doing this, for every dataset we compute the significance of the noise-line outlier using sky shifting through the five-vector method (Sec. II C 2). As above, we produce 10^4 background realizations using sky patches separated with a minimal distance of 0.01 deg. For comparison with the traditional method, we additionally evaluate the significance using the theoretical formula assuming pure-Gaussian noise (cf. Fig. 4) [46].

We evaluate our method’s ability to identify noise artifacts by studying the rejection probability as a function of the confidence threshold set to claim a detection—that is, how likely the analysis is to reject the artifact as we decrease our tolerance for false alarms (FAP). Figure 10 shows the results obtained empirically with sky shifting (right panel) and analytically assuming Gaussian noise (left panel). In the ideal case, we would be able to perfectly measure the significance of an outlier and the curves in Fig. 10 would simply be a straight line with slope -1 . As we can see, strong ($\text{SNR} > 256$) noise lines produce significant outliers if we assume the background to be Gaussian; however, this is not true for sky shifting.

These tests can be extended to a general noise background. In principle, we can model coherent instrumental noise as a linear combination of monochromatic noise lines like those injected above. Every noise line will couple constructively or destructively with the other noise lines after Doppler corrections. If the lines combined

constructively, we would obtain a case very similar to the one presented in Fig. 10 but with a larger noise-line SNR. On the other hand, if the noise lines combine destructively, then the SNR would be lower than Fig. 10. The general case should lie somewhere in between.

V. CONCLUSION

Sky-shifting can provide a much-needed efficient and robust way to empirically estimate the background of searches for continuous gravitational waves targeted at known pulsars, enabling estimates of detection significance that are valid in actual (non-Gaussian) instrumental noise. This method has already been put into practice for diagnosing outliers in actual LIGO and Virgo searches [16,24], but a systematic study of its performance was lacking from the literature. In this paper, we fill in this gap by introducing the rationale behind this strategy, exploring its theoretical applicability and studying its performance in real and simulated data.

The procedure is simple: The original gravitational-wave data are time corrected for multiple shifted sky locations to obtain as many instantiations of noise-only data, which are then analyzed by any of the usual searches with the same settings as the on-source search (Sec. III A). Under the right conditions, we show that the sky-shifted data are blind to astrophysical signals while retaining the statistical properties of the noise. This allows for the direct empirical estimation of the background distribution of the different search statistics.

Two conditions need to be satisfied for sky-shifting to be effective: Shifted sky locations must (i) be resolvably different and (ii) be drawn from the same hemisphere as the source. As long as this is true, sky shifting will provide

independent draws from the background distribution (Sec. III B). Furthermore, for realistic signal amplitudes, the distribution will be uncontaminated by the presence of a signal at the true on-source sky location. This is not true for extremely loud signals, but this is not a problem because in those cases the on-source statistic is always louder than the background (Sec. III B 2). The phenomenon is analogous to that observed with strong signals in searches for compact binaries [61,62].

We illustrate the efficacy of sky shifting with several examples in real and synthetic data (Sec. IV A). This includes simulated Gaussian noise in the absence of signal (Fig. 4), as well as in the presence of realistic (network SNR = 5, Fig. 5) and strong (network SNR = 700, Fig. 6) signals. We also demonstrate the method in the presence of real LIGO O1 noise with data prepared for PSR J1932 + 17 [16] and loud hardware injection [64]. Source parameters for all these case studies are summarized in Table I.

Finally, we systematically study the performance of sky shifting by looking at false-dismissal and false-alarm rates (Sec. IV B). The former is quantified by the receiver operating curve of Fig. 9 and the latter by the rejection-probability vs confidence-level plot of Fig. 10. In the cases we consider, we find that sky shifting outperforms the standard method for computing significances in the context of the five-vector search.

ACKNOWLEDGMENTS

We thank Jonah Kanner and Alan J. Weinstein for insightful comments and suggestions regarding the method; we thank Thomas Dent for useful feedback, including suggesting the “sky-shift” nomenclature; we also thank Graham Woan for reviewing this manuscript, as well as colleagues in the LIGO Scientific Collaboration Continuous Waves group for useful discussions. M.I. is supported by NASA through the NASA Hubble Fellowship Grant No. HST-HF2-51410.001-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under Contract No. NAS5-26555. LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science Foundation and operates under Cooperative Agreement No. PHY-0757058. We are grateful for computational resources provided by Cardiff University, and funded by a STFC grant supporting UK Involvement in the Operation of Advanced LIGO. For the period of this work M.P. was funded by the STFC under Grant No. ST/N005422/1. S.M. thanks the LIGO Visitor Program. Plots produced using Matplotlib [67]. This paper carries LIGO Document No. LIGO-P2000413.

-
- [1] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Observation of Gravitational Waves from a Binary Black Hole Merger, *Phys. Rev. Lett.* **116**, 061102 (2016).
 - [2] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral, *Phys. Rev. Lett.* **119**, 161101 (2017).
 - [3] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs, *Phys. Rev. X* **9**, 031040 (2019).
 - [4] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GW190425: Observation of a compact binary coalescence with total mass $\sim 3.4 M_{\odot}$, *Astrophys. J. Lett.* **892**, L3 (2020).
 - [5] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GW190412: Observation of a binary-black-hole coalescence with asymmetric masses, *Phys. Rev. D* **102**, 043015 (2020).
 - [6] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GW190814: Gravitational waves from the coalescence of a 23 solar mass black hole with a 2.6 solar mass compact object, *Astrophys. J.* **896**, L44 (2020).
 - [7] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GW190521: A Binary Black Hole Merger with a Total Mass of $150 M_{\odot}$, *Phys. Rev. Lett.* **125**, 101102 (2020).
 - [8] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Properties and astrophysical implications of the $150 M_{\odot}$ binary black hole merger GW190521, *Astrophys. J. Lett.* **900**, L13 (2020).
 - [9] J. Aasi *et al.* (LIGO Scientific Collaboration), Advanced LIGO, *Classical Quantum Gravity* **32**, 115012 (2015).
 - [10] F. Acernese *et al.* (Virgo Collaboration), Advanced Virgo: A second-generation interferometric gravitational wave detector, *Classical Quantum Gravity* **32**, 024001 (2015).
 - [11] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaboration), Results of the deepest all-sky survey for continuous gravitational waves on LIGO S6 data running on the Einstein@Home volunteer distributed computing project, *Phys. Rev. D* **94**, 102002 (2016).
 - [12] J. Aasi *et al.* (LIGO Scientific and Virgo Collaborations), Searches for continuous gravitational waves from nine young supernova remnants, *Astrophys. J.* **813**, 39 (2015).
 - [13] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Comprehensive all-sky search for periodic

- gravitational waves in the sixth science run LIGO data, *Phys. Rev. D* **94**, 042002 (2016).
- [14] J. Aasi *et al.* (LIGO Scientific and Virgo Collaborations), First all-sky search for continuous gravitational waves from unknown sources in binary systems, *Phys. Rev. D* **90**, 062010 (2014).
- [15] J. Aasi *et al.* (LIGO Scientific and Virgo Collaborations), First low frequency all-sky search for continuous gravitational wave signals, *Phys. Rev. D* **93**, 042007 (2016).
- [16] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), First search for gravitational waves from known pulsars with advanced LIGO, *Astrophys. J.* **839**, 12 (2017).
- [17] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Searches for gravitational waves from known pulsars at two harmonics in 2015–2017 LIGO data, *Astrophys. J.* **879**, 10 (2019).
- [18] K. S. Thorne, in *Three Hundred Years of Gravitation*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1987), Chap. 9, pp. 330–458.
- [19] J. Aasi *et al.* (LIGO Scientific and Virgo Collaborations), Gravitational waves from known pulsars: Results from the initial detector era, *Astrophys. J.* **785**, 119 (2014).
- [20] R. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Gravitational-wave constraints on the equatorial ellipticity of millisecond pulsars, *Astrophys. J. Lett.* **902**, L21 (2020).
- [21] K. Riles, Recent searches for continuous gravitational waves, *Mod. Phys. Lett. A* **32**, 1730035 (2017).
- [22] M. Pitkin, M. Isi, J. Veitch, and G. Woan, A nested sampling code for targeted searches for continuous gravitational waves from pulsars, [arXiv:1705.08978](https://arxiv.org/abs/1705.08978).
- [23] P. Jaranowski, A. Królak, and B. F. Schutz, Data analysis of gravitational-wave signals from spinning neutron stars: The signal and its detection, *Phys. Rev. D* **58**, 063001 (1998).
- [24] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), First low-frequency Einstein@Home all-sky search for continuous gravitational waves in advanced LIGO data, *Phys. Rev. D* **96**, 122004 (2017).
- [25] S. A. Usman *et al.*, The PyCBC search for gravitational waves from compact binary coalescence, *Classical Quantum Gravity* **33**, 215004 (2016).
- [26] T. Dal Canton *et al.*, Implementing a search for aligned-spin neutron star-black hole systems with advanced ground based gravitational wave detectors, *Phys. Rev. D* **90**, 082004 (2014).
- [27] N. J. Cornish and L. Sampson, Towards robust gravitational wave detection with pulsar timing arrays, *Phys. Rev. D* **93**, 104047 (2016).
- [28] D. I. Jones and N. Andersson, Gravitational waves from freely precessing neutron stars, *Mon. Not. R. Astron. Soc.* **331**, 203 (2002).
- [29] R. T. Edwards, G. B. Hobbs, and R. N. Manchester, TEMPO2, a new pulsar timing package—II. The timing model and precision estimates, *Mon. Not. R. Astron. Soc.* **372**, 1549 (2006).
- [30] R. Dupuis and G. Woan, Bayesian estimation of pulsar parameters from gravitational wave data, *Phys. Rev. D* **72**, 102002 (2005).
- [31] L. Sammut, C. Messenger, A. Melatos, and B. Owen, Implementation of the frequency-modulated sideband search method for gravitational waves from low mass x-ray binaries, *Phys. Rev. D* **89**, 043001 (2014).
- [32] P. Astone, A. Colla, S. D’Antonio, S. Frasca, and C. Palomba, Method for all-sky searches of continuous gravitational wave signals using the frequency-Hough transform, *Phys. Rev. D* **90**, 042002 (2014).
- [33] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GW150914: The Advanced LIGO Detectors in the Era of First Discoveries, *Phys. Rev. Lett.* **116**, 131103 (2016).
- [34] D. V. Martynov *et al.*, Sensitivity of the advanced LIGO detectors at the beginning of gravitational wave astronomy, *Phys. Rev. D* **93**, 112004 (2016); Erratum, *Phys. Rev. D* **97**, 059901 (2018).
- [35] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Effects of data quality vetoes on a search for compact binary coalescences in advanced LIGOs first observing run, *Classical Quantum Gravity* **35**, 065010 (2018).
- [36] P. Covas *et al.* (LSC Instrument Authors), Identification and mitigation of narrow spectral artifacts that degrade searches for persistent gravitational waves in the first two observing runs of advanced LIGO, *Phys. Rev. D* **97**, 082002 (2018).
- [37] P. B. Covas, A. Effler, E. Goetz, P. M. Meyers, A. Neunzert, M. Oliver, B. L. Pearlstone, V. J. Roma, R. M. S. Schofield, V. B. Adya *et al.*, Identification and mitigation of narrow spectral artifacts that degrade searches for persistent gravitational waves in the first two observing runs of advanced LIGO, *Phys. Rev. D* **97**, 082002 (2018).
- [38] P. Leaci, P. Astone, M. A. Papa, and S. Frasca, Using a cleaning technique for the search of continuous gravitational waves in LIGO data, *J. Phys. Conf. Ser.* **228**, 012006 (2010).
- [39] G. D. Meadors, K. Kawabe, and K. Riles, Increasing LIGO sensitivity by feedforward subtraction of auxiliary length control noise, *Classical Quantum Gravity* **31**, 105014 (2014).
- [40] J. Driggers *et al.* (LIGO Scientific Collaboration), Improving astrophysical parameter estimation via offline noise subtraction for advanced LIGO, *Phys. Rev. D* **99**, 042001 (2019).
- [41] T. M. Niebauer, A. Rüdiger, R. Schilling, L. Schnupp, W. Winkler, and K. Danzmann, Pulsar search using data compression with the Garching gravitational wave detector, *Phys. Rev. D* **47**, 3106 (1993).
- [42] M. Pitkin, Searches for continuous and transient gravitational waves from known neutron stars and their astrophysical implications, Ph.D. Thesis, University of Glasgow, 2006.
- [43] P. Jaranowski and A. Królak, Searching for gravitational waves from known pulsars using the \mathcal{F} and \mathcal{G} statistics, *Classical Quantum Gravity* **27**, 194015 (2010).
- [44] D. Keitel, R. Prix, M. A. Papa, P. Leaci, and M. Siddiqi, Search for continuous gravitational waves: Improving robustness versus instrumental artifacts, *Phys. Rev. D* **89**, 064023 (2014).
- [45] D. Keitel and R. Prix, Line-robust statistics for continuous gravitational waves: Safety in the case of unequal detector sensitivities, *Classical Quantum Gravity* **32**, 035004 (2015).

- [46] P. Astone, S. D’Antonio, S. Frasca, and C. Palomba, A method for detection of known sources of continuous gravitational wave signals in non-stationary data, *Classical Quantum Gravity* **27**, 194016 (2010).
- [47] P. Astone, A. Colla, S. D’Antonio, S. Frasca, and C. Palomba, Coherent search of continuous gravitational wave signals: Extension of the 5-vectors method to a network of detectors, *J. Phys. Conf. Ser.* **363**, 012038 (2012).
- [48] P. Astone, A. Colla, S. D’Antonio, S. Frasca, C. Palomba, and R. Serafinelli, Method for narrow-band search of continuous gravitational wave signals, *Phys. Rev. D* **89**, 062008 (2014).
- [49] D. Keitel, R. Prix, M. A. Papa, P. Leaci, and M. Siddiqi, Search for continuous gravitational waves: Improving robustness versus instrumental artifacts, *Phys. Rev. D* **89**, 064023 (2014).
- [50] M. Isi, M. Pitkin, and A. J. Weinstein, Probing dynamical gravity with the polarization of continuous gravitational waves, *Phys. Rev. D* **96**, 042001 (2017).
- [51] S. Mastrogiovanni, P. Astone, S. D’Antonio, S. Frasca, G. Intini, P. Leaci, A. Miller, C. Palomba, O. J. Piccinni, and A. Singhal, An improved algorithm for narrow-band searches of continuous gravitational waves, *Classical Quantum Gravity* **34**, 135007 (2017).
- [52] P. Astone, A. Colla, S. D’Antonio, S. Frasca, C. Palomba, and R. Serafinelli, Method for narrow-band search of continuous gravitational wave signals, *Phys. Rev. D* **89**, 062008 (2014).
- [53] O. J. Piccinni, P. Astone, S. D’Antonio, S. Frasca, G. Intini, P. Leaci, S. Mastrogiovanni, A. Miller, C. Palomba, and A. Singhal, A new data analysis framework for the search of continuous gravitational wave signals, *Classical Quantum Gravity* **36**, 015008 (2019).
- [54] R. Prix and J. T. Whelan, \mathcal{F} -statistic search for white-dwarf binaries in the first mock LISA data challenge, *Classical Quantum Gravity* **24**, S565 (2007).
- [55] D. Keppel, Multidetector F -statistic metric for short-duration nonprecessing inspiral gravitational-wave signals, *Phys. Rev. D* **86**, 123010 (2012).
- [56] S. Mastrogiovanni *et al.*, Phase decomposition of the template metric for continuous gravitational-wave searches, *Phys. Rev. D* **98**, 102003 (2018).
- [57] M. Isi, A. J. Weinstein, C. Mead, and M. Pitkin, Detecting beyond-Einstein polarizations of continuous gravitational waves, *Phys. Rev. D* **91**, 082002 (2015).
- [58] K. Wette and R. Prix, Flat parameter-space metric for all-sky searches for gravitational-wave pulsars, *Phys. Rev. D* **88**, 123005 (2013).
- [59] K. Wette, Parameter-space metric for all-sky semicoherent searches for gravitational-wave pulsars, *Phys. Rev. D* **92**, 082003 (2015).
- [60] S. Mastrogiovanni, P. Astone, S. D. Antonio, S. Frasca, G. Intini, I. La Rosa, P. Leaci, A. Miller, F. Muciaccia, C. Palomba *et al.*, Phase decomposition of the template metric for continuous gravitational-wave searches, *Phys. Rev. D* **98**, 102003 (2018).
- [61] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Binary Black Hole Mergers in the First Advanced LIGO Observing Run, *Phys. Rev. X* **6**, 041015 (2016).
- [62] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), GW150914: First results from the search for binary black hole coalescence with advanced LIGO, *Phys. Rev. D* **93**, 122003 (2016).
- [63] J. Abadie *et al.* (LIGO Scientific and Virgo Collaborations), Search for gravitational waves from low mass compact binary coalescence in LIGO’s sixth science run and Virgo’s science runs 2 and 3, *Phys. Rev. D* **85**, 082002 (2012).
- [64] C. Biwer *et al.*, Validating gravitational-wave detections: The advanced LIGO hardware injection system, *Phys. Rev. D* **95**, 062002 (2017).
- [65] B. P. Abbott *et al.* (Virgo and LIGO Scientific Collaboration), First Search for Nontensorial Gravitational Waves from Known Pulsars, *Phys. Rev. Lett.* **120**, 031104 (2018).
- [66] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), First narrow-band search for continuous gravitational waves from known pulsars in advanced detector data, *Phys. Rev. D* **96**, 122006 (2017).
- [67] J. D. Hunter, Matplotlib: A 2D graphics environment, *Comput. Sci. Eng.* **9**, 90 (2007).