5*n*-vector ensemble method for detecting gravitational waves from known pulsars

Luca D'Onofrio,[1,2,*] Rosario De Rosa,[1,2] Luciano Errico,[1,2] Cristiano Palomba,[3] Valeria Sequino,[1,2] and Lucia Trozzo[2]

[1]*Universit di Napoli Federico II, I-80126 Napoli, Italy*
[2]*INFN, Sezione di Napoli, I-80126 Napoli, Italy*
[3]*INFN, Sezione di Roma, I-00185 Roma, Italy*

(Received 13 January 2022; accepted 25 February 2022; published 15 March 2022)

We present a multiple test for the targeted search of continuous gravitational waves from an ensemble of known pulsars, combining multidetector single pulsar statistics defined through the 5*n*-vector method. In order to maximize the detection probability, we describe a rank truncation method to select the most promising sources within the ensemble, based on the p-values computed for single pulsar analysis. To test the efficiency of our method, we use a Monte Carlo procedure and define a p-value for the ensemble that is an overall p-value for the hypothesis of continuous wave emissions from a set of known pulsars. We also perform a pilot search on real data from the O3A dataset of the two LIGO detectors.

DOI: 10.1103/PhysRevD.105.063012

I. INTRODUCTION

Continuous gravitational waves (CWs) are promising targets for the interferometric gravitational waves detectors. CWs are persistent and quasimonochromatic signals that can be originated from nonaxisymmetric spinning neutron stars, and/or potentially from other exotic sources [1].

The standard definition for the gravitational wave amplitude for a nonprecessing triaxial star rotating around a principal axis with rotation frequency f_{rot} is [2]

$$h_0 = \frac{16\pi^2 G}{c^4}\frac{I\epsilon f_{\mathrm{rot}}^2}{d}, \tag{1}$$

where I is the star moment of inertia with respect to the rotation axis, d is the distance from the Earth and ϵ, the stars ellipticity.

So far, no CW signal has been detected in the analysis of interferometric detectors data. Recent searches from the three observing runs (O1, O2, and O3A) for the LIGO and Virgo observatories [3–7] set stringent constraints on neutron star ellipticity.

In the latest results for targeted search [7], where the source parameters (sky position, frequency, and frequency derivatives) are assumed to be known with high accuracy, the LIGO and Virgo Collaborations set direct observational limits on the amplitude that are, for the first time, two millisecond pulsars-close (for J0437-4715) or below (for J0711-6830) the theoretical spin-down limit. This allows us to constrain the fraction of rotational energy which may be lost due to the gravitational-wave emissions.

Our work tries to improve the detection probability for the targeted search of CWs, considering an ensemble of individually undetectable pulsars. Specifically, in this paper we describe a procedure to test the hypothesis that CWs are emitted from a selected set of pulsars.

Traditional multiple tests consider either each hypothesis separately, as in the Bonferroni-like procedures [8], or all hypotheses simultaneously as in the case of the Fisher test [9] (or other p-value combinations [10]). In order to improve the detection probability, since only few signals are expected to dominate over the majority, it may be more desirable to consider the combined evidence for subsets of the hypotheses, as in the case of [11,12].

In CW analysis, the first method to detect CWs from an ensemble of known pulsars was proposed by Giazotto *et al.* in [13]. Recently, other methods have been presented by Cutler and Schutz [14] and by Fan *et al.* [15] combining F-statistic values, by Pitkin *et al.* using a Hierarchical Bayesian method [16], and by Buono *et al.* in [17] combining 5-vector statistics values [18].

In this paper, we generalize the method described in [17], defining a multidetector ensemble statistic using the 5*n*-vector method (the extension of the 5-vector to a network of detectors [19]). Since we expect that only a small number of signals could be detectable with the current detectors' sensitivity, we propose a rank truncation method that improves the detection probability controlling the look-elsewhere effect [20]. We test the sensitivity of the method using a Monte Carlo procedure and considering different prior assumptions on the

*ldonofrio@na.infn.it

strength of the expected signals. We also test the method considering real data and a set of simulated pulsars.

The paper is organized as follows. In Sec. II we study the ensemble statistic as the linear combination of single pulsar statistics, considering different choices for the coefficients of the combination. In Sec. III we describe the rank-truncation method that allows us to select the top ranking sources according to the significance of the corresponding individual test. In Sec. IV, using a Monte Carlo algorithm, we infer the noise distribution of the proposed statistic as a function of the number of pulsars considered in the ensemble. In Sec. V we describe a Monte Carlo procedure to estimate the sensitivity of the method and a possible strategy for determining upper limits. In Sec. VI, we generalize the pipeline considering real data from the O3A dataset of the two LIGO detectors. Conclusions are presented in Sec. VII. In the Appendixes, we describe some mathematical aspects of the proposed method.

II. THE DETECTION STATISTIC

The ensemble statistic can be defined as the weighted linear combination of the N multidetector single pulsar statistics,

$$t = \sum_{i=1}^{N} (b_{+,i} S_{+,i} + b_{\times,i} S_{\times,i}), \tag{2}$$

where the $S_{+/\times,i}$ are defined through the $5n$-vector method (with n the number of considered detectors, see Appendix A). A theoretical choice for the coefficients $b_{+/\times,i}$ can be obtained by maximizing the critical ratio (see Appendix B),

$$\bar{b}_{+/\times,i} = \frac{|H_{+/\times,i}|^2 |\mathbf{A}_i^{+/\times}|^8 H_{0,i}^2}{(\sum_{j=1}^{n} \sigma_j^2 \cdot T_j \cdot |\mathbf{A}_{j,i}^{+/\times}|^2)^2}, \tag{3}$$

where σ_j^2, T_j are the variance of the Gaussian data distribution around the signal frequency (usually one tenth of a Hz wide) and the observation time in the jth detector, respectively. $|\mathbf{A}_i^{+/\times}|^2 = \sum_{j=1}^{n} |\mathbf{A}_{j,i}^{+/\times}|^2$ where $|\mathbf{A}_{j,i}^{+/\times}|^2$ is the 5-vector signal plus/cross template for the ith pulsar in the jth detector. The coefficients in (3) depend also on the polarization functions $H_{+/\times,i}$ and on the amplitude $H_{0,i}$ that are unknown in a real analysis.

The ensemble statistic with coefficients in Eq. (3) is the theoretical limit that we want to approach with an appropriate choice for the coefficients. In [17] a joint ensemble statistic using 5-vectors and combining multidetector statistics for single pulsar was proposed, defined as the weighted sum of the detection statistics in each detector.

We propose a new ensemble statistic, based on the $5n$-vectors, with coefficients,

$$b_{+/\times,i} = \frac{|\mathbf{A}_i^{+/\times}|^4}{\sum_{j=1}^{n} \sigma_j^2 \cdot T_j \cdot |\mathbf{A}_{j,i}^{+/\times}|^2}. \tag{4}$$

As described in Appendix C, the coefficients in Eq. (4) normalize the single-pulsar statistic with respect to the detectors' sensitivity and observation time. In this way, the noise distribution of this normalized single pulsar statistic is the same for each pulsar, while the signal distribution also has an analytic expression in contrast to [17,19].

Figure 1 compares the different definitions considering a set of simulated signals (see Appendix D). We ranked these signals by decreasing values of α, defined as the ratio between the injected amplitude H and the minimum detectable amplitude $h_{\min}$ [21]

$$H = \alpha \cdot h_{\min} \approx \alpha \cdot 11 \sqrt{\frac{S(f_{gw})}{T_{\text{obs}}}}, \tag{5}$$

where $S(f_{gw})$ is the one-sided power spectral density at the expected signal frequency f_{gw}, and T_{obs} is the observation time.

In Fig. 1, we considered two ideal detectors with sensitivity and observation time equal to the LIGO Livingston (LLO) design case for the O3A run. As α prior distribution with $0.01 < \alpha < 0.6$, we chose an exponential distribution with mean value equal to 0.09.

The continuous line, obtained with coefficients in Eq. (4), approaches the t definition with coefficients in Eq. (3) and outperforms the t definition in [17]. In Fig. 1, the detection probability increases twofold compared to single detector case and is 20% better compared to [17].

For each definition (except for the dashed line where the coefficients are amplitude-weighted), the detection

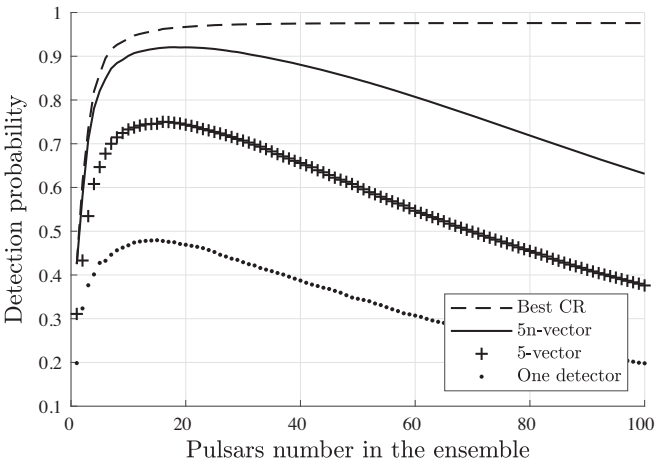

FIG. 1. Detection probability for a fixed false alarm probability of 1% increasing the number of simulated pulsars in the ensemble and considering Gaussian noise in two detectors equal to LLO in O3A run. The injected signals are ranked by decreasing values of α [see Eq. (5)]. The dashed line is obtained considering the t definition with coefficients in (3), the continuous line with coefficients in (4), the '+' and '·' line with the t definition using 5-vectors as in [17]. The '·' line is obtained considering only one detector.

probability increases with an increasing number of signals in the ensemble up to a maximum and then starts to decrease. This is linked to the prior exponential distribution that fixes the signals' strength; by adding smaller and smaller signals we do not contribute to the ensemble signal but, rather, to the noise.

III. THE RANK TRUNCATION METHOD

As shown in Fig. 1, in order to improve the detection probability for the ensemble method, we need to rank pulsars trying to estimate each signal strength. Indeed, in a real analysis we do not know the α parameter that ranks the sources in Fig. 1.

In [17], we used the single pulsar p-value as a statistical parameter to rank the sources in the ensemble. Nevertheless, by increasing the number of pulsars, a series of nonsignificant results may together suggest significance. This is the well-known look-elsewhere effect.

Moreover, ranking pulsars for increasing p-values means ranking pulsars for decreasing values of the normalized single pulsar statistic [defined in brackets in Eq. (2) with coefficients in Eq. (4),

$$\bar{S}_{(1)} < \bar{S}_{(2)} < \ldots < \bar{S}_{(N)}. \tag{6}$$

By considering the measured values $\bar{S}_{(i)}$ as order statistics (see Appendix E), we can control the look-elsewhere effect since the $\bar{S}_{(i)}$ distribution depends on N, the number of measured statistics (i.e., the number of analyzed pulsars).

We propose the partial sum $T(k)$ of the ordered statistics values,

$$T(k) = \sum_{i=N-k+1}^{N} \bar{S}_{(i)} \tag{7}$$

as the ensemble detection statistic for our rank truncation method. Indeed, the $T(k)$ distribution depends on N for each k; in this way we combine pulsars with the smallest p-values that are assumed near the detection threshold controlling the look-elsewhere effect.

If we knew that among the N pulsars there is at most one signal, the optimal procedure to control the look-elsewhere effect would be the Šidák correction, as used in CWs narrow-band analysis [4,22].

For $k = N$, the distribution of $T(N)$ coincides with the distribution of t.

For $k = 1$, $T(1)$ is the largest order statistic since $T(1) = \bar{S}_{(N)}$, with known analytic distribution (see Appendix E and Fig. 2 for the noise case).

In the general case $1 < k < N$, the convolution of order statistics has no simple expression. The complexity of the analytic form of $T(k)$ is due to the dependency introduced by ordering the p-values; when the $(k+1)$th p-value, that is a random variable, happens to be relatively small, the k smallest p-values have to squeeze into a relatively tiny interval from zero to that value. As shown in Sec. IV and Sec. V, we use a Monte Carlo algorithm to recover the $T(k)$ distributions as a function of k.

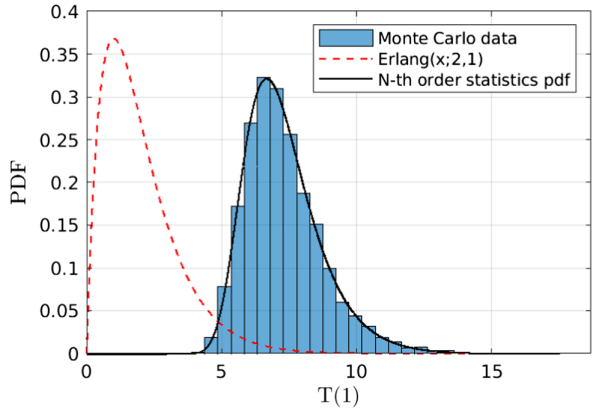

FIG. 2. Noise distribution (blue histogram) for $T(1)$ defined in Eq. (7) for $k = 1$ and inferred using the Monte Carlo algorithm described in Sec. IV. The dashed line is the single pulsar noise distribution using the normalized statistic. The continuous line is the theoretical probability distribution function for the largest order statistic for $N = 100$.

IV. SUM OF ORDER STATISTICS DISTRIBUTION

In this section we describe the Monte Carlo algorithm used to determine the $T(k)$ noise distribution. We fixed $N = 100$; that is, an ensemble of 100 pulsars.

First, we have considered the case of Gaussian noise with zero mean value. Therefore, the normalized single pulsar statistic, with coefficients given in Eq. (4), have a common Erlang$(x; 2, 1)$ distribution, that is a Gamma distribution with positive integer shape and scale parameters (see Appendix C).

To infer the $T(k)$ noise distribution, we have used the following Monte Carlo algorithm.

(1) Generate an Erlang$(x; 2, 1)$ distribution with 200,000 points to simulate single-pulsar noise distribution.
(2) Select randomly N points to simulate an ensemble detection.
(3) Rank these N points for decreasing values (that is for increasing p-values).
(4) Repeat steps 1–3, 10,000 times.

The variable $T(k)$ is the sum of the k largest points selected each times. Empirically, we have found that a gamma function can fit the distribution to a good approximation for all values of k. As shown in Fig. 3, the distribution of $T(N)$ coincides with the distribution of t, since the shape parameter approaches $2N = 200$ and the scale parameter tends towards 1, as expected. Indeed, the sum of Erlang random variables with the same scale parameter follows an

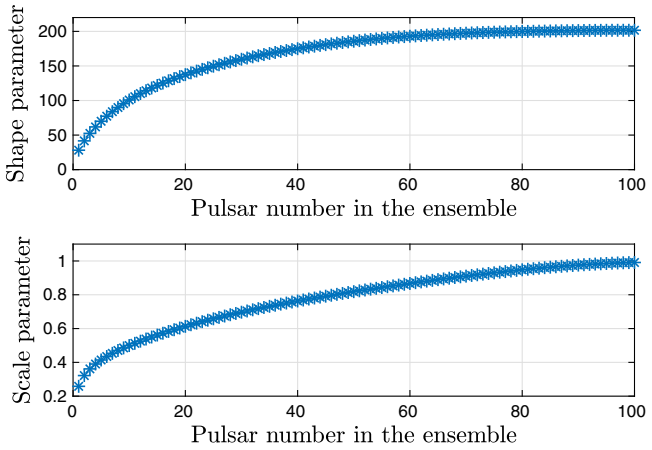

FIG. 3. Shape and scale parameters as a function of k, inferred from the fit to the $T(k)$ noise distributions using a Gamma distribution.

Erlang distribution with shape parameter equal to the sum of the shape parameters [i.e., Erlang$(x; 2N, 1)$].

The fitting parameters in Fig. 3 depend only on N, the number of considered pulsars, and on the assumption of Gaussian noise for the detectors' data.

V. P-VALUE FOR THE ENSEMBLE

In this section we test our method for different choices of the prior assumption on the expected signals' strength. By considering a set of N sources, we compute a p-value for the overall hypothesis that all single hypotheses—or part of them—are true using the statistic $T(k)$. The signal distribution for the normalized single-pulsar statistic is linked to a noncentral χ^2 distribution with noncentrality parameter Λ defined in Eq. (C5) for $N = 1$. To fix the signal strength, we have used different prior distributions for the Λ parameter. Specifically, we have considered an exponential distribution with different mean values $\bar{\Lambda}$ and a flat distribution with only ten signals, fixing the Λ value to have a given P_d, i.e., the detection probability for single pulsar at a fixed false alarm probability of 1%.

To test our method, and also to study the sensitivity of the ensemble analysis, we have applied the following Monte Carlo procedure.

(1) Fix the prior distribution for Λ.
(2) Generate N noncentral χ^2 distributions with 200,000 points with the appropriate Λ to simulate individual signal distributions.
(3) Select randomly N points, one from each distribution to simulate an ensemble detection.
(4) Rank these N points for decreasing values (that is, for increasing p-values).
(5) Compute the p-value for the ensemble, as a function of k, by using the parameters from the gamma fit to the $T(k)$ noise distribution.
(6) Repeat steps 2–5, 10,000 times.

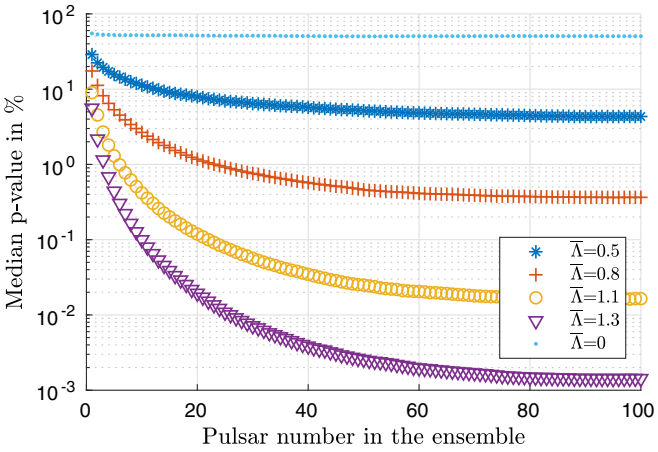

FIG. 4. Median for the p-value distribution computed using the Monte Carlo algorithm described in Sec. V as a function of k and for different mean values $\bar{\Lambda}$ of the Λ prior exponential distribution. In any case, the detection probability for single pulsar is at most 40%. For $\bar{\Lambda} = 0$, Λ is fixed to zero for all pulsars.

(7) For each k, compute the median value for the 10,000 p-values of ensemble.

As shown in Fig. 4, if $\Lambda = 0$ for each pulsar, the median of the p-value for $T(k)$ is almost 50% since in the noise hypothesis, the p-value follows a uniform distribution for each k. As expected, the median of the p-value decreases by increasing the mean value $\bar{\Lambda}$.

The set of median p-values is an indication of the sensitivity of the method since you have 50% of probability to obtain a set of p-values below the considered curve for the fixed $\bar{\Lambda}$.

Figure 5 shows that using $T(k)$, we improved the detection probability even if there are few signals in the

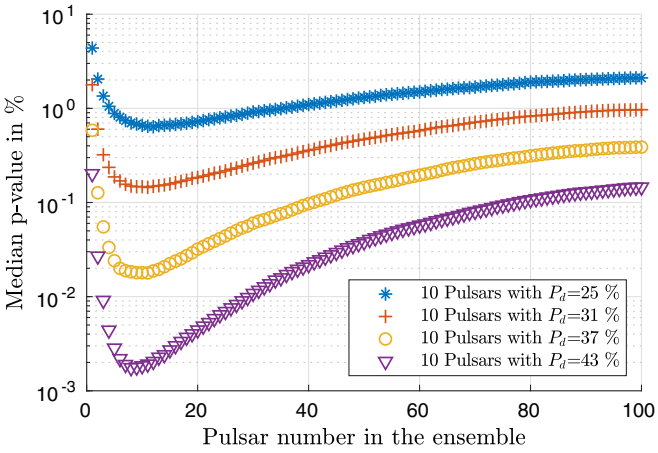

FIG. 5. Median for the p-value distribution computed using the Monte Carlo algorithm described in Sec. V as a function of k. Λ is different from zero for ten pulsars and is fixed considering P_d, the detection probability for single pulsar at the fixed false alarm probability of 1%.

ensemble. If in the analyzed ensemble there are ten signals with $P_d = 25\%$, there is almost a 50% probability of obtaining a p-value for the ensemble below the 1% false alarm threshold.

It is clear that fixing the number of signals and P_d for each signal, you can improve the detection probability for decreasing N. In a real analysis, however, we cannot optimize N and we need a criteria to select the most promising sources. This criteria should be based on the data introducing an order in the set of pulsars. This is, for example, the case of the coherence, an independent parameter that measures the resemblance between the shape of the expected signal and the data [18].

The coherence is not strictly related to the value of the measured statistic but it can happen that the highest values for the coherence correspond to the smallest p-values, also in the hypothesis of no signal.

We suggest to fix N according to physical observations, for example considering pulsars with upper limit on the strain amplitude below the spin-down limit or considering the set of millisecond pulsars.

A. Upper limit

In the case of no evidence for CW signals from the ensemble, we can consider all the selected pulsars to compute the analytic signal distribution for the $T(k)$ statistic. Indeed, the $T(N)$ distribution coincides with the t distribution obtained in Appendix C.

The aim is to provide an upper limit on the set of amplitudes by using the t signal distribution.

According to Eq. (C5), Λ is a function of the unknown N-squared amplitudes and of the unknown pulsar polarization parameters, η and ψ. It can be rewritten as

$$\Lambda = \sum_{i=1}^{N} H_{0,i}^2 \cdot f_i(\eta, \psi). \tag{8}$$

f_i depends also on the sky position and on the detector sensitivity at the expected frequency for the ith pulsar, that are known parameters for a targeted search.

Marginalizing over the unknown polarization parameters, assuming a uniform distribution, we can define an approximated value $\overline{f_i}$ for each $f_i(\eta, \psi)$. We have verified that, with this approximation, the error for each f_i is below 5%.

It follows that,

$$\Lambda \sim \sum_{i=1}^{N} H_{0,i}^2 \cdot \overline{f_i}. \tag{9}$$

From the measured value of the ensemble statistic, we can estimate the value $\Lambda^{95\%}$ that entails a detection probability of 95% according to the analytic signal distribution. Considering Eq. (9) and $\Lambda^{95\%}$, we can obtain an upper limit on the sum of the squared amplitudes.

VI. APPLICATION TO O3A DATA

In this section we describe the application of the ensemble search method to real data, considering LIGO Livingston and LIGO Hanford O3A run datasets [23]. To test the method, we injected fake signals into the O3A data considering the set of pulsars in Appendix D, and fixed appropriate amplitudes that matched the desired Λ.

For single pulsar analysis, we used the Band Sample Data (BSD) framework, as in [7]. The BSD framework [24] is based on the construction of BSD files, i.e., complex time series, each covering 10 Hz and 1 month of the original dataset. Using the BSD files, the computational cost of the full-coherent analysis is reduced to a few CPU-minutes per source per detector.

Since noise in real data does not necessarily follow a Gaussian distribution, we constructed the experimental distributions for the normalized single pulsar statistic considering a range of off-source frequencies as in [22]. Figure 6 shows the fitted Gamma parameters to the experimental distributions of the normalized statistic for each pulsar. For several pulsars, the fitted shape and scale parameters tend towards 2 and 1, respectively, as expected for Gaussian noise distribution.

To generalize the Monte Carlo algorithm described in Sec. IV, we replaced the Erlang distribution in step one with the N different experimental noise distributions inferred from the data. We constructed the $T(k)$ statistics, and for each k the fitted shape and scale parameters are shown in Fig. 7 and compared with the parameters in Fig. 3 for the Gaussian noise case.

As shown in Sec. V, the sensitivity of the method depends mainly on the fixed distribution of Λ and on the number N of pulsars in the ensemble. Considering real data, the sensitivity depends also on the GW frequencies that determine the frequency bands where the signals are expected.

Figure 8 shows that the results obtained in Sec. V are also valid in the case of real data. Indeed, as also shown in

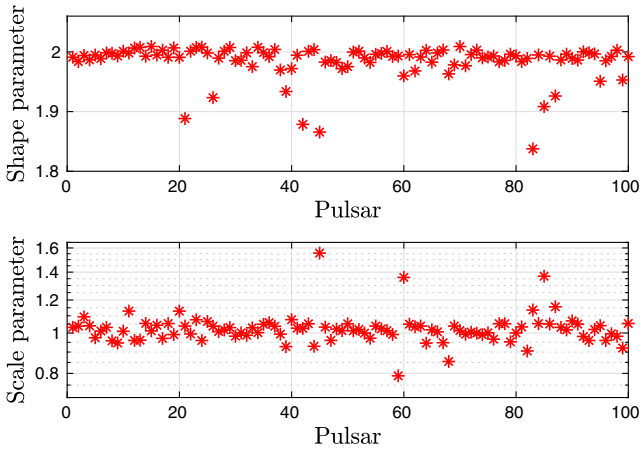

FIG. 6. Gamma parameters inferred from the Gamma fit to the experimental distributions of the normalized single pulsar statistics for each pulsar.

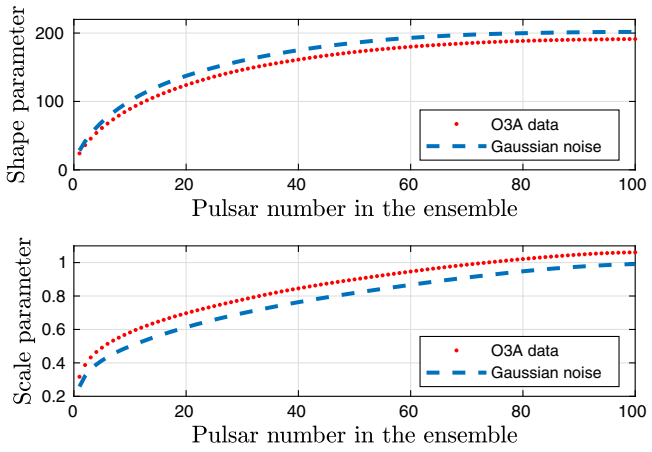

FIG. 7. Shape and scale parameters inferred from the Gamma fit to $T(k)$ as a function of k in the case of Gaussian noise as in the Fig. 3 (dashed line) and in the case of O3A data (dot line).

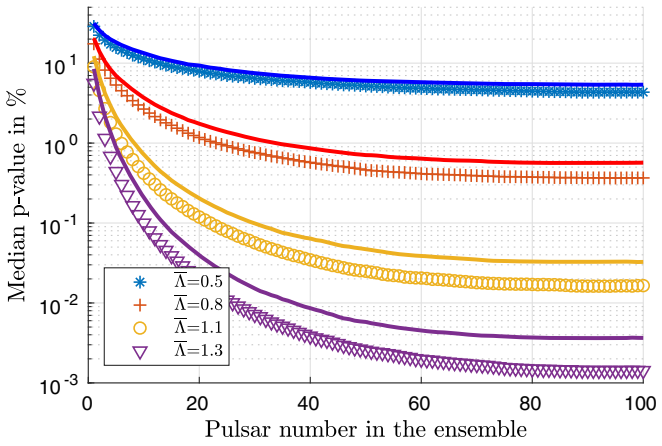

FIG. 8. Median for the p-value distribution computed using the Monte Carlo algorithm described in Sec. V as a function of k and for different mean value $\bar{\Lambda}$ of the Λ prior exponential distribution. The continuous lines are inferred from O3A data and compared to the results of Fig. 4.

Fig. 6, the Gaussian distribution for the noise is a good approximation for the O3A dataset.

Since when the noise distribution is not perfectly Gaussian the tails of the statistic distributions may be heavier, the sensitivity of the ensemble method slightly decreases compared to Fig. 4. For the set of analyzed pulsars and O3A data, we have more than a 50% probability to obtain a p-value of the ensemble below the 1% detection threshold for an exponential Λ distribution with mean value of 0.8.

VII. CONCLUSION

In this work, we propose a multidetector ensemble method for detecting CWs signals from a set of individually undetectable pulsars.

Since few signals are expected near the detection threshold, we optimize the detection probability for the ensemble by considering a subset of the total number of analyzed pulsars. The procedure is based on a rank truncation method, since we consider the top ranked statistics to detect the combined effects of a few individually undetectable signals.

The proposed ensemble statistic $T(k)$ combines the k top ranked statistics for single pulsar p-value, defined through the $5n$-vector method. The procedure can be easily generalized to different pipelines, for example using the F-statistics.

To study the performances of the method, we constructed the distribution for the $T(k)$ p-value as a function of k, considering different signal strengths for the combined individual tests. As shown in Sec. V, the ensemble method improves the detection probability for CW signals. In case of no detection, we can also provide an upper limits on the sum of the squared amplitudes for the analyzed pulsars.

The results, described in Fig. 3 and in Fig. 4, depend on the number N of the analyzed pulsars, and on the power of the combined individual tests and on the Gaussian noise assumption.

To apply the ensemble procedure to real data, we generalize the results in Fig. 2 for the $T(k)$ noise distributions, combining the experimental noise distributions for each pulsar statistic inferred from the data.

In Fig. 8, we show that using the O3A dataset for the two LIGO detectors the sensitivity of the method slightly decreases compared to the Gaussian noise case.

We plan to apply the ensemble method to the most recent LIGO and Virgo observation runs and to the set of pulsars analyzed in targeted searches. The potential detection of a gravitational wave emission from an ensemble cannot provide information on the single-pulsar parameters, but it would give clear evidence of the presence of CW sources in the Galaxy.

ACKNOWLEDGMENTS

This research has made use of data obtained from the Gravitational Wave Open Science Center [25], a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. LIGO Laboratory and Advanced LIGO are funded by the United States National Science Foundation (NSF) as well as the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Planck-Society (MPS), and the State of Niedersachsen/Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. Virgo is funded, through the European Gravitational Observatory (EGO), by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale di Fisica Nucleare

(INFN) and the Dutch Nikhef, with contributions by institutions from Belgium, Germany, Greece, Hungary, Ireland, Japan, Monaco, Poland, Portugal, Spain.

APPENDIX A: 5n-VECTOR METHOD

The 5n-vectors [19],

$$\mathbf{X} = [\mathbf{X}_1, \ldots, \mathbf{X}_n], \qquad \mathbf{A}^+ = [\mathbf{A}_1^+, \ldots, \mathbf{A}_n^+],$$
$$\mathbf{A}^\times = [\mathbf{A}_1^\times, \ldots, \mathbf{A}_n^\times] \tag{A1}$$

combine the data 5-vectors $\mathbf{X}_j$ and the template 5-vectors $\mathbf{A}_j^{+/x}$ ([18], with $j = 1, \ldots, n$) for the considered pulsar in the n detectors.

Using the 5n-vectors, the multidetector single pulsar statistic S is defined as

$$S = |\mathbf{A}^+|^4 |\hat{H}_+|^2 + |\mathbf{A}^\times|^4 |\hat{H}_\times|^2, \tag{A2}$$

where (the same for $\hat{H}_\times$)

$$\hat{H}_+ = \frac{\mathbf{X}\cdot\mathbf{A}^+}{|\mathbf{A}^+|^2} = \frac{\sum_{j=1}^n \mathbf{X}_j \cdot (\mathbf{A}_j^+)^*}{\sum_{k=1}^n \mathbf{A}_k^+ \cdot (\mathbf{A}_k^+)^*}$$
$$= \frac{1}{|\mathbf{A}^+|^2}(|\mathbf{A}_1^+|^2 \cdot \hat{H}_{+,1} + \cdots + |\mathbf{A}_n^+|^2 \cdot \hat{H}_{+,n}). \tag{A3}$$

Each data 5-vector interacts only with the corresponding template.

In the hypothesis of Gaussian noise with zero mean and variance σ_j^2 in each detector, the components of the data 5-vector are also distributed according to a complex Gaussian distribution with mean value zero and variance $\sigma_j^2 \cdot T_j$ where T_j is the observation time for the jth detector. The two complex estimators $\hat{H}_{+/\times}$ have also Gaussian distributions,

$$\hat{H}_{+/\times} \sim \mathrm{Gauss}(x; 0, \sigma_{+/\times}^2) \tag{A4}$$

with

$$\sigma_{+/\times}^2 = \sum_{j=1}^n \frac{\sigma_j^2 \cdot T_j \cdot |\mathbf{A}_j^{+/\times}|^2}{|\mathbf{A}^{+/\times}|^4}. \tag{A5}$$

Since $|\hat{H}_{+/\times}|^2 = \mathrm{Re}[\hat{H}_{+/\times}]^2 + \mathrm{Im}[\hat{H}_{+/\times}]^2$, it follows that

$$|\hat{H}_{+/\times}|^2 \sim \mathrm{Exp}(x; \sigma_{+/\times}^2) = \frac{1}{\sigma_{+/\times}^2} e^{-\frac{x}{\sigma_{+/\times}^2}}. \tag{A6}$$

We can compute the noise S distribution considering the weighted linear combination in (A2).

If a CW signal is present in the analyzed data, the distributions of the two complex estimators $\hat{H}_{+/\times}$ are

$$\hat{H}_{+/\times} \sim \mathrm{Gauss}(x; H_0 \cdot e^{j\gamma} \cdot H_{+/\times}, \sigma_{+/\times}^2), \tag{A7}$$

where $H_{+/\times}$ are the polarization functions, H_0 is the amplitude and γ the phase.

$|\hat{H}_{+/\times}|^2$ have a noncentral-χ^2 distribution (apart from the factor $k_{+/\times}$)

$$|\hat{H}_{+/\times}|^2 \sim \frac{k_{+/\times}}{2} e^{-\frac{k_{+/\times}x + \lambda_{+/\times}}{2}} I_0(\sqrt{k_{+/\times}\lambda_{+/\times}x}), \tag{A8}$$

where I_0 is the modified Bessel function of the first kind, and

$$k_{+/\times} = \frac{2}{\sigma_{+/\times}^2}, \qquad \lambda_{+/\times} = k_{+/\times} \cdot |H_{+/\times}|^2 \cdot H_0^2. \tag{A9}$$

Compared to the single-detector case, the distributions are the same. The difference is the expression for the variances $\sigma_{+/\times}^2$.

APPENDIX B: THEORETICAL COEFFICIENTS

Let us consider a random variable $Y = \sum_{i=1}^N a_i X_i$ that is the linear combination of N random variables X_i with unknown coefficients a_i. The critical ratio (CR) is defined as

$$\mathrm{CR} = \frac{(\mu_{\mathrm{sig}} - \mu_n)^2}{\Theta_n^2}, \tag{B1}$$

where μ_{sig} is the mean of the signal distribution of Y, while μ_n and Θ_n^2 are, respectively, the mean and variance of Y when there is noise only. If the signal distributions of X_i are noncentral-χ^2 distributions with K degrees of freedom and noncentrality parameter Λ_i, the CR is

$$\mathrm{CR} = \frac{(\sum_{i=1}^N a_i \Lambda_i)^2}{2K \sum_{k=1}^N a_k^2}. \tag{B2}$$

The coefficients that maximize the CR are

$$\bar{a}_i = \Lambda_i. \tag{B3}$$

In the case of the F-statistic, the distribution is a $4D$ χ^2 distribution with a noncentrality parameter equal to the squared optimal signal to noise ratio ρ^2 in the case of signal [2]. Linearly combining F-statistic values from different pulsars, the coefficients that maximize the CR are $\bar{a}_i = \rho_i^2$, in agreement with [15].

In [17], the 5-vector ensemble statistic t can be written as

$$t = \sum_{i=1}^N a_i S_i = \sum_{i=1}^N (b_{+,i} S_{+,i} + b_{\times,i} S_{\times,i}), \tag{B4}$$

where N is the pulsars number in the ensemble and $S_{+/\times}$ have 2-D χ^2 distributions [apart from the factor $k_{+/\times}$; see Eqs. (A6) and (A8)]. Therefore, the coefficients that maximize the CR are

$$\bar{b}_{+/\times,i} = \lambda_{+/\times,i} \cdot k_{+/\times,i} = \frac{|H_{+/\times,i}|^2 |\mathbf{A}_i^{+/\times}|^4 H_{0,i}^2}{\sigma_i^4 \cdot T_{\text{obs}}^2}. \quad \text{(B5)}$$

These results are in agreement with [17], since the coefficients that maximize the CR for the single-pulsar ($N = 1$) statistic S are $b_{+/\times} = |H_{+/\times}|^2 |\mathbf{A}^{+/\times}|^4$.

If S is defined using the $5n$-vector method, the coefficients of t that maximize the CR are

$$\bar{b}_{+/\times,i} = \lambda_{+/\times,i} \cdot k_{+/\times,i} = \frac{|H_{+/\times,i}|^2 |\mathbf{A}_i^{+/\times}|^8 H_{0,i}^2}{(\sum_{j=1}^n \sigma_j^2 \cdot T_j \cdot |\mathbf{A}_{j,i}^{+/\times}|^2)^2}, \quad \text{(B6)}$$

where $|\mathbf{A}_i^{+/\times}|^2 = \sum_{j=1}^n |\mathbf{A}_{j,i}^{+/\times}|^2$ and T_j is the observation time in the jth detector. To check this, if $n = 1$ the expression in (B6) is equal to (B5).

APPENDIX C: ENSEMBLE STATISTIC DISTRIBUTIONS

The coefficients in (4) are equal to the inverse of the variance in (A4)

$$b_{+/\times,i} = \frac{|\mathbf{A}_i^{+/\times}|^4}{\sum_{j=1}^n \sigma_j^2 \cdot T_j \cdot |\mathbf{A}_{j,i}^{+/\times}|^2} \equiv \frac{1}{\sigma_{+/\times,i}^2}. \quad \text{(C1)}$$

In the case of Gaussian noise and considering (A6), the distribution of $b_{+/\times,i} \cdot S_{+/\times,i}$ is

$$b_{+/\times,i} \cdot S_{+/\times,i} \equiv b_{+/\times,i} \cdot |\hat{H}_{+/\times,i}|^2 \sim \text{Exp}(x;1) = e^{-x} \quad \text{(C2)}$$

The ensemble statistic t in (2) is the sum of $2N$ exponential random variables with mean values equal to one. It follows that

$$t \sim \text{Erlang}(x; 2N, 1) = \frac{x^{(2N-1)} e^{-x}}{(2N-1)!}, \quad \text{(C3)}$$

which is the Erlang distribution with the scale parameter equal to one and the shape parameter equal to twice the number of pulsars N.

If a signal is present into the data, $b_{+/\times,i} \cdot S_{+/\times,i}$ distribution is proportional to noncentral-χ^2 distribution,

$$\begin{aligned} b_{+/\times,i} \cdot S_{+/\times,i} &\sim 2 \cdot \text{Chi}(2x; 2, \lambda_{+/\times}) \\ &= e^{-x + \frac{\lambda_{+/\times}}{2}} I_0(\sqrt{\lambda_{+/\times} x}). \end{aligned} \quad \text{(C4)}$$

Therefore, t is the sum of $2N$ noncentral-χ^2 random variables. The signal distribution is the noncentral-χ^2 distribution with $4N$ degrees of freedom and noncentrality parameter Λ,

$$\begin{aligned} \Lambda &= \sum_{i=1}^N (\lambda_{+,i} + \lambda_{\times,i}) = \sum_{i=1}^N H_{0,i}^2 (k_{+,i} \cdot |H_{+,i}|^2 + k_{\times,i} \cdot |H_{\times,i}|^2) \\ &= \sum_{i=1}^N 2 \cdot H_{0,i}^2 \left(\frac{|\mathbf{A}_i^+|^4 \cdot |H_{+,i}|^2}{\sum_{j=1}^n \sigma_j^2 \cdot T_j \cdot |\mathbf{A}_{j,i}^+|^2} + \frac{|\mathbf{A}_i^\times|^4 \cdot |H_{\times,i}|^2}{\sum_{k=1}^n \sigma_k^2 \cdot T_k \cdot |\mathbf{A}_{k,i}^\times|^2} \right). \end{aligned} \quad \text{(C5)}$$

APPENDIX D: SIMULATED PULSARS

To test the pipeline, we used a set of simulated pulsars. For each pulsar, we randomly fixed the sky position. We chose a uniform distribution for the polarization parameters, an exponential distribution for the GW frequency between 20 Hz and 120 Hz (see Fig. 9), and a uniform distribution for the first derivative of the GW frequency. We fixed the sensitivity and the observation time considering the O3A design sensitivity for LIGO Livingston and LIGO Hanford.

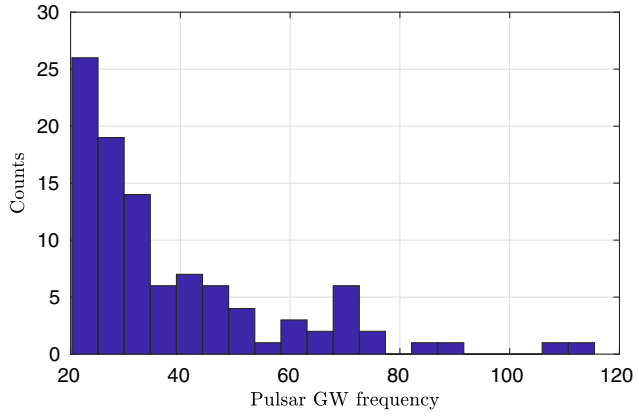

FIG. 9. Histogram of the GW frequencies in Hz for the set of simulated pulsars used to test the ensemble method.

APPENDIX E: ORDER STATISTICS

Let $X_1, X_2, \ldots, X_N$ be a set of independent and identically distributed random variables. Let $F(x)$ and $f(x)$ be the cumulative distribution function (cdf) and the probability distribution function (pdf), respectively. Consider a realization/measurement for these N random variables: the k-th order statistic $X_{(k)}$ is defined as the kth smallest value of the obtained sample [26],

$$X_{(1)} < X_{(2)} < \ldots < X_{(k)} < \ldots < X_{(N)}, \quad \text{(E1)}$$

$X_{(k)}$ takes the value of X_k if X_k is the kth random variable when the realizations are ranked in ascending order. It is straightforward to infer the cumulative distribution function $F_N(x)$ for $X_{(N)}$

$$F_N(x) = P(X_{(N)} < x) = P(X_1 < x \cup \ldots \cup X_N < x) = P(X_1 < x) \cup \ldots \cup P(X_N < x) = [F(x)]^N. \quad \text{(E2)}$$

It follows that the probability distribution function $f_N(x)$ is

$$f_N(x) = \frac{dF_N(x)}{dx} = N \cdot [F(x)]^{(N-1)} \cdot f(x) \quad \text{(E3)}$$

With similar considerations we can infer that the cumulative distribution function and the probability distribution function for $X_{(1)}$ are

$$F_1(x) = P(X_{(1)} < x) = 1 - P(X_{(1)} > x) = 1 - [P(X_1 > x) \cup \ldots \cup P(X_N > x)] = 1 - [1 - F(x)]^N, \quad \text{(E4)}$$

$$f_1(x) = \frac{dF_1(x)}{dx} = N \cdot [1 - F(x)]^{(N-1)} \cdot f(x). \quad \text{(E5)}$$

The smallest- and largest order statistic /pdf are special cases of the kth order statistic cdf/pdf,

$$F_k(x) = P(X_{(k)} < x) = \sum_{i=k}^{N} \binom{N}{i} [F(x)]^i [F(x)]^{N-i}, \quad \text{(E6)}$$

$$f_k(x) = \frac{N!}{(k-1)!(N-K)!} [F(x)]^{k-1} [1 - F(x)]^{N-k} f(x). \quad \text{(E7)}$$

[1] K. Glampedakis and L. Gualtieri, Gravitational waves from single neutron stars: An advanced detector era survey, in *The Physics and Astrophysics of Neutron Stars*, edited by L. Rezzolla, P. Pizzochero, D. I. Jones, N. Rea, and I. Vidaña (Springer International Publishing, Cham, 2018).

[2] 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).

[3] B. Abbott *et al.* (LIGO Scientific Collaboration and the Virgo Collaboration), Searches for gravitational waves from known pulsars at two harmonics in 2015–2017 LIGO data, Astrophys. J. **879**, 10 (2019).

[4] B. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Narrow-band search for gravitational waves from known pulsars using the second LIGO observing run, Phys. Rev. D **99**, 122002 (2019).

[5] B. Abbott *et al.* (LIGO Scientific Collaboration and the Virgo Collaboration), All-sky search in early O3 LIGO data for continuous gravitational-wave signals from unknown neutron stars in binary systems, Phys. Rev. D **103**, 064017 (2021).

[6] B. Abbott *et al.* (LIGO Scientific Collaboration and the Virgo Collaboration), Diving below the Spin-down Limit: Constraints on Gravitational Waves from the Energetic Young Pulsar PSR J0537-6910, Astrophys. J. Lett. **913**, L27 (2021).

[7] R. Abbott *et al.* (LIGO Scientific Collaboration and the Virgo Collaboration), Gravitational-wave constraints on the equatorial ellipticity of millisecond pulsars, Astrophys. J. Lett. **902**, L21 (2020).

[8] S. Chen, Z. Feng, and X. Yi, A general introduction to adjustment for multiple comparisons, J. Thorac. Dis. **9** (2017).

[9] R. Fisher, Statistical methods for research workers, in *Breakthroughs in Statistics: Methodology and Distribution*, edited by S. Kotz and N. L. Johnson (Springer, , New York, 1992).

[10] S. Won, N. Morris, Q. Lu, and R. C. Elston, Choosing an optimal method to combine p-values, Stat. Med. **28** (2009).

[11] D. Zaykin *et al.*, Truncated product method for combining p-values, Genet. Epidemiol. **22**, 170 (2002).

[12] F. Dudbridge and B. Koeleman, Rank truncated product of p-values, with application to genomewide association scans, Genet. Epidemiol. **25**, 360 (2003).

[13] A. Giazotto, S. Bonazzola, and E. Gourgoulhon, Gravitational waves emitted by an ensemble of rotating neutron stars, Phys. Rev. D **55**, 2014 (1997).

[14] C. Cutler and B. F. Schutz, Generalized $\mathcal{F}$-statistic: Multiple detectors and multiple gravitational wave pulsars, Phys. Rev. D **72**, 063006 (2005).

[15] X. Fan, Y. Chen, and C. Messenger, Method to detect gravitational waves from an ensemble of known pulsars, Phys. Rev. D **94**, 084029 (2016).

[16] M. Pitkin, C. Messenger, and X. Fan, Hierarchical Bayesian method for detecting continuous gravitational waves from an ensemble of pulsars, Phys. Rev. D **98**, 063001 (2018).

[17] M. Buono, R. De Rosa, L. D'Onofrio, L. Errico, C. Palomba, O. Piccinni, and V. Sequino, A method for detecting continuous gravitational wave signals from an ensemble of known pulsars, Classical Quantum Gravity **38** (2021).

[18] 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).

[19] 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** (2012).

[20] R. G. J. Miller, *Simultaneous Statistical Inference* (Springer, New York, 1981).

[21] P. Leaci, Searching for continuous gravitational wave signals using LIGO and Virgo detectors, J. Phys. Conf. Ser. **354** (2012).
[22] 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** (2014).
[23] R. Abbott *et al.*, Open data from the first and second observing runs of Advanced LIGO and Advanced Virgo, SoftwareX **13**, 100658 (2021).
[24] O. J. Piccinni, P. Astone, S. D'Antonio, S. Frasca, G. Intini, P. Leaci, S. Mastrogiovanni, A. L. 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).
[25] https://www.gw-openscience.org/.
[26] H. A. David and H. N. Nagaraja, *Order Statistics* (Wiley Series in Probability and Statistics, New York 2003).