# Novel directed search strategy to detect continuous gravitational waves from neutron stars in low- and high-eccentricity binary systems

Paola Leaci,[1,2,*] Pia Astone,[2] Sabrina D'Antonio,[3] Sergio Frasca,[1,2] Cristiano Palomba,[2] Ornella Piccinni,[1,2] and Simone Mastrogiovanni[1,2]

[1]*Dipartimento di Fisica, Università di Roma "Sapienza", Piazzale Aldo Moro, 2, I-00185 Rome, Italy*
[2]*INFN, Sezione di Roma, Piazzale Aldo Moro, 2, I-00185 Rome, Italy*
[3]*INFN, Sezione di Roma 2, Via della Ricerca Scientifica, 1, I-00133 Roma, Italy*

We describe a novel, very fast and robust, directed search incoherent method (which means that the phase information is lost) for periodic gravitational waves from neutron stars in binary systems. As a directed search, we assume the source sky position to be known with enough accuracy, but all other parameters (including orbital ones) are supposed to be unknown. We exploit the frequency modulation due to source orbital motion to unveil the signal signature by commencing from a collection of time and frequency peaks (the so-called "peakmap"). We validate our algorithm (pipeline), adding 131 artificial continuous-wave signals from pulsars in binary systems to simulated detector Gaussian noise, characterized by a power spectral density $S_h = 4 \times 10^{-24}$ Hz$^{-1/2}$ in the frequency interval [70, 200] Hz, which is overall commensurate with the advanced detector design sensitivities. The pipeline detected 128 signals, and the weakest signal injected (added) and detected has a gravitational-wave strain amplitude of $\sim 10^{-24}$, assuming one month of gapless data collected by a single advanced detector. We also provide sensitivity estimations, which show that, for a single-detector data covering one month of observation time, depending on the source orbital Doppler modulation, we can detect signals with an amplitude of $\sim 7 \times 10^{-25}$. By using three detectors, and one year of data, we would easily gain a factor 3 in sensitivity, translating into being able to detect weaker signals. We also discuss the parameter estimate proficiency of our method, as well as computational budget: sifting one month of single-detector data and 131 Hz-wide frequency range takes roughly 2.4 CPU hours. Hence, the current procedure can be readily applied in ally-sky schemes, sieving in parallel as many sky positions as permitted by the available computational power. Finally, we introduce (ongoing and future) approaches to attain sensitivity improvements and better accuracy on parameter estimates in view of the use on real advanced detector data.

## I. INTRODUCTION

Since the early 1960s, when the first gravitational-wave bar detector was developed [1], the experiments that aimed at the detection of gravitational radiation, planned in laboratories throughout the world, have been in continuous progress [2–10]. At present, LIGO and Virgo [11–13] are the most sensitive ground-based gravitational-wave detectors.

Following a major upgrade that lasted for five years, with consequent improvement in sensitivity [14], the LIGO observatories resumed data taking with the first observing science run during which they have collected data between September 2015 and January 2016. On September 14, 2015, the advanced LIGO interferometers detected for the first time a coincident transient gravitational-wave signal produced by the coalescence of a pair of black holes [15], marking thus the official beginning of a new era: the era of gravitational-wave astronomy.

There are, however, other classes of gravitational-wave signals, which have still to be detected, such as long-lived continuous waves (CWs), which are expected to be emitted by rapidly rotating neutron stars (NSs) with nonaxisymmetric deformations [16]. Based on standard theory of stellar evolution, we expect $\mathcal{O}(10^8)$ of these sources to exist in the Galaxy [17], but only $\sim$2500 NSs (mostly pulsars) have been electromagnetically observed [18]. Roughly 1300 of these observed radio pulsars are located in binary systems, and have rotation rates that can allegedly emit CWs in the advanced LIGO-Virgo sensitivity band. This promising class of sources is the target of the current paper.

The detection of CW signals will enrich the understanding that we have about the emitting objects (i.e., NSs), providing us insight about the equation of state of the matter at supranuclear densities inside the NSs [19], know exactly the NS degree of asymmetry, and have also demographic and evolutionary information about these sources.

In the rest frame of the NS, CWs have a constant amplitude and are quasimonochromatic with a slowly decreasing intrinsic frequency. They are received at Earth-based detectors with a Doppler modulation due to the relative motion between the source and the detector. Consequently, the observed phase evolution depends on the

*paola.leaci@roma1.infn.it

intrinsic signal frequency, frequency time derivatives (also known as "spindown" terms), and source sky position. If the source is located in a binary system, there is a further frequency-modulation caused by the source orbital motion, which in general is described by five unknown Keplerian parameters [20] (as detailed in Sec. II B).

The weakness of the expected signal requires long integration times, typically of the order of a few months or years, to accumulate a signal-to-noise ratio (SNR) sufficient for detection as, for a coherent (incoherent.[1]) search, the SNR scales as the square (fourth) root of the length of the observational time (i.e., the length of the data being analyzed) [21,22]. All-sky, wide frequency searches over long observation times cannot be treated by using standard coherent methods (where the phase information is used), as is the case for targeted and narrow-band searches for known pulsars [23,24], because of the demanding computational burden. Hence, hierarchical approaches have been proposed [25–27], where the entire data set is split into shorter segments. Each segment is analyzed coherently, and afterwards the information from the different segments is combined incoherently. The hierarchical approaches allow us to dramatically reduce the analysis computational time at the cost of a relatively small sensitivity loss.

The additional source orbital parameters make the search parameter space to blow up, resulting in a prohibitive computational cost [28]. Hence, it becomes pressing to develop robust strategies to detect CWs emitted by NSs orbiting a companion object, and being able to reach a tradeoff between computational cost and sensitivity. Although CWs have not been detected so far by analyzing data from initial LIGO and Virgo detectors, stringent upper limits have been set on the gravitational-wave signal strength for both isolated pulsars [24,29–31] and pulsars in binary systems [32,33].

A particularly interesting type of potential CW sources are NSs in low-mass x-ray binaries (LMXBs), with Scorpius X-1 being its most prominent representative [34]. Several searches for CW signals from Scorpius X-1 have been performed (without any detections) on data from initial LIGO [33,35], and new algorithms (pipelines) have been developed [28] and recently tested in a Scorpius X-1 mock data challenge (MDC) [36].

We present here an incoherent strategy that allows us to perform directed searches for CWs emitted by NSs in binary systems (LMXB like sources), exploring a wide frequency range and source orbital parameters by using only a standard laptop computer for a few hours. We also include investigations of pulsars in eccentric orbits, which we know to exist [18,37,38]. The method is based on selecting significant peaks from a short fast-Fourier-transform (FFT) database (SFDB), and exploiting the

frequency modulation pattern produced by the source orbital motion to detect a potential CW signal. We show the performance of the current method to detect CW signals by analyzing pure Gaussian noise data to which we add hundreds of fake signals. We consider one month of gapless data, i.e., data taken continuously during the assumed observation time from the Virgo (or equivalently LIGO) detector in its advanced configuration.

We restrict our investigation to constant-frequency CW signals (i.e., we neglect frequency time derivatives). This is motivated by the assumed steady-state torque-balance situation in LMXB like sources [39,40], which are our main target of interest.[2] However, the corresponding fluctuations in the accretion rate are expected to cause some stochastic frequency drift, and one will therefore need to be careful to restrict the maximal *coherence time* (i.e., FFT duration) in order to limit the frequency resolution. In fact, in [28] it has been shown that the maximal FFT duration would be restricted by the astrophysical concern of *spin wandering*, namely a stochastic variability of the spin frequency due to variations in the accretion rate. In the present work we neglect the spin wandering effect, as our new robust methodology is expected to be unaffected by these variations. Exhaustive future studies will be able to shed light on these considerations and will be presented in a subsequent paper.

The novel procedure described here allows us to detect gravitational-wave signals with strain amplitude of $\sim 10^{-24}$, the weakest value that we used for the simulations, assuming one month of gapless data collected by a single advanced detector. Sensitivity estimation studies show however that, depending on the source orbital Doppler modulation, the current method can detect signals with an amplitude of $\sim 7 \times 10^{-25}$. By using three advanced detectors, and one year of data, we would be able to further improve the sensitivity by a factor 3. This novel algorithm is the first in the literature able to provide estimates for orbital eccentricity and argument of periapse, and this is achieved with a small computational cost.

The paper is organized as follows. Section II provides a general and brief introduction of the expected signal model. Section III describes the data analysis approach. In Secs. V and IV, we present a rigorous description of the novel

---

[1]In an incoherent method the phase information is lost.

[2]The plausibility of the torque-balance scenario is based on the observed lack of accreting neutron stars spinning at very high frequencies (above 700 Hz) [41], even though their ages and accretion rates imply that they should have accreted enough matter to reach the centrifugal break-up limit around 1400 Hz [42]. Alternative interpretations of the lack of high spin frequencies in the most rapidly accreting neutron stars have been proposed and involve disc accretion physics [43]. Although these interpretations indicate that gravitational radiation is not required to brake the spin-up of the neutron star, they do not rule out gravitational emission from accreting systems. We underline that the torque-balance argument is used here as an approximate bound.

strategy used to detect CW signals orbiting a companion object. In Sec. VI, we report the parameter estimation methods. In Sec. VII, we discuss the choices of the investigated parameter space, and detail the signal generation. Sections VIII and IX show the key results of the procedure. Sensitivity estimates and computing cost budget are detailed in Secs. X and XI, respectively. Finally, Sec. XII contains concluding remarks, underway search improvements, and future prospects.

## II. THE SIGNAL

In the following we briefly recall the expected waveform model and signal phase.

### A. The waveform model

The expected waveform of a nonaxisymmetric NS, rapidly rotating around one of its principal axes, and received at the detector is

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

where $\iota$ is the inclination angle of the NS rotational axis to the line of sight and $\phi(t)$ is the phase evolution of the gravitational-wave signal; $F_{+,\times}$ are the detector beam pattern functions to *plus* and *cross* polarized gravitational waves, which depend on the sky position $\vec{n}$ and the relative polarization angle $\psi$ of the wave-frame[3] [21,44]. In standard equatorial coordinates, with right ascension $\alpha$ and declination $\delta$, the components of the unit vector $\vec{n}$, pointing to the NS, are given by $(\cos \alpha \cos \delta, \sin \alpha \cos \delta, \sin \delta)$.

The gravitational-wave amplitude parameter is given by

$$h_0 = \frac{4\pi^2 G}{c^4} \frac{I_{zz} f^2 \varepsilon}{d}, \quad (2)$$

where $f$ is the frequency of the emitted CW signal (which is also twice the rotational frequency of the star for the sources we are interested in [45]); $G$ is Newton's constant, $c$ is the speed of light, $d$ is the distance to the star, and $\varepsilon$ is the star ellipticity expressed in terms of principal moments of inertia. The distribution of $\varepsilon$ for NSs is uncertain and model dependent since the breaking strain for a NS crust is highly uncertain (see, e.g., [46–49] for exhaustive discussion).

Spinning NSs in binary systems are particularly interesting because accretion from a companion may cause an asymmetrical quadrupole moment of inertia of the spinning

NS. An intriguing astrophysical model postulating torque balance between the accretion and CW emission [39,40,50] yields a predicted CW amplitude that, for Scorpius X-1, is

$$h_0 \sim 3.5 \times 10^{-26} \sqrt{\frac{300 \text{ Hz}}{f}}. \quad (3)$$

### B. Binary CW signal phase

The general CW phase model assumes a slowly spinning-down NS with a rotation rate and quadrupolar deformation resulting in the emission of CWs. The phase evolution can therefore be expressed as a Taylor series in the NS source frame as

$$\phi^{\text{src}}(\tau) = 2\pi[f(\tau - t_{\text{ref}}) + \frac{1}{2}\dot{f}(\tau - t_{\text{ref}})^2 + \cdots], \quad (4)$$

where $t_{\text{ref}}$ denotes the reference time and $f, \dot{f}, \ddot{f}, \ldots$ are the CW frequency and spindown parameters.

To relate the CW phase in the source frame to the phase $\phi(t_{\text{arr}})$ in the detector frame, we need to relate the wavefront detector arrival time $t_{\text{arr}}$ to its source emission time $\tau$, i.e., $\tau(t_{\text{arr}})$, such that $\phi(t_{\text{arr}}) = \phi^{\text{src}}(\tau(t_{\text{arr}}))$. Neglecting relativistic wave-propagation effects, such as Einstein and Shapiro delays (see, e.g., [51,52] for more details), we can write this as

$$\tau(t_{\text{arr}}) = t_{\text{arr}} + \frac{\vec{r}(t_{\text{arr}}) \cdot \vec{n}}{c} - \frac{D}{c} - \frac{R(\tau)}{c}, \quad (5)$$

where $\vec{r}$ is the vector from solar-system barycenter (SSB) to the detector, $D$ is the (generally unknown) distance between the SSB and the binary barycenter (BB), $R$ is the radial distance of the CW-emitting NS from the BB along the line of sight, where $R > 0$ means the NS is further away from us than the BB.

Following the discussion in Sec. III A of [28], we can write the Rømer delay of the binary (i.e., the light-travel time across the orbit) as

$$\frac{R}{c} = a_p[\sin \omega(\cos E - e) + \cos \omega \sin E \sqrt{1 - e^2}], \quad (6)$$

where we defined the projected semimajor axis $a_p \equiv a \sin I / c$ of the NS orbit; $I$ is the inclination angle between the orbital plane and the sky, $a$ the semimajor axis, $\omega$ is the argument of periapse, $e$ the orbital eccentricity, and $E$ the eccentric anomaly, defined by the transcendental relation

$$\tau - t_p = \frac{P}{2\pi}(E - e \sin E). \quad (7)$$

Equation (7) is the so-called Kepler's equation, and describes the dynamics in binary systems; $P$ is the binary period, and $t_p$ is the time of periapse passage.

---

[3]The wave frame is a right-handed Cartesian coordinate system based on the direction of propagation of the gravitational wave. Its $z$ axis is along the direction of propagation, and its x-and y-axes are along the principal directions of polarization of the wave.

Dropping the unknown distance $D$ to the BB (which is equivalent to redefining the intrinsic spindown parameters), and defining the SSB wavefront arrival time $t_{SSB}$ as

$$t_{SSB}(t_{arr}; \vec{n}) \equiv t_{arr} + \frac{\vec{r}(t_{arr}) \cdot \vec{n}}{c}, \qquad (8)$$

we can rewrite the timing relation Eq. (5) as

$$\tau(t_{SSB}) = t_{SSB} - \frac{R(\tau)}{c}. \qquad (9)$$

As we are interested only in binary systems with known sky-position $\vec{n}$, we can place us into the SSB, which is always possible for known $\vec{n}$ [28]. In order to simplify the notation, we now simply write $t \equiv t_{SSB}$. Plugging this timing model into the phase of Eq. (4), we obtain

$$\phi(t) \approx 2\pi \left[ f\left(\Delta t - \frac{R}{c}\right) + \frac{1}{2}\dot{f}\left(\Delta t - \frac{R}{c}\right)^2 + \cdots \right], \quad (10)$$

with $\Delta t \equiv t - t_{ref}$. The binary systems we are interested in have semimajor axis $a_p$ of order of $\mathcal{O}(1\text{–}3)$ s, and binary periods $P$ of order of several hours. Hence, the change in $E$, and therefore $R(E)$ during the time $R/c$, will be negligible, and so we can approximate $E(\tau) \approx E(t)$, namely

$$t - t_p \approx \frac{P}{2\pi}(E - e \sin E). \qquad (11)$$

For our purposes, by using the linear phase-model approximation, we can write Eq. (10) as [28]

$$\phi(t) \approx 2\pi f\left[\Delta t - \frac{R(t)}{c}\right]. \qquad (12)$$

Replacing Eq. (6) into Eq. (12), we find

$$\phi(t) = 2\pi f\{\Delta t - a_p[\sin\omega(\cos E(t) - e)$$
$$+ \cos\omega \sin E(t)\sqrt{1 - e^2}]\}, \qquad (13)$$

which is the phase model valid for eccentric orbits. It is useful to consider, however, also the small-eccentricity limit,[4] thus simplifying Eqs. (6) and (13). To this purpose, we Taylor expand Eqs. (6) and (11) up to leading order in $e$, i.e., inserting $E(t) = E_0(t) + eE_1(t) + \cdots$ into Kepler's equation [i.e., Eq. (11)], and obtain

$$E_0(t) = \Omega(t - t_p), \qquad (14)$$

$$E_1(t) = \sin E_0(t). \qquad (15)$$

Plugging this information into Eq. (6), we obtain the Rømer delay of the binary to leading order in $e$ as

-----

[4]Low-eccentricity values generally range from 0 up to 0.1 [28].

$$\frac{R(t)}{c} = a_p\left[\sin\psi(t) + \frac{\kappa}{2}\sin 2\psi(t) - \frac{\eta}{2}\cos 2\psi(t) - \frac{3\eta}{2}\right]. \qquad (16)$$

We use the standard Laplace-Lagrange parameters defined as

$$\kappa \equiv e\cos(\omega), \qquad (17)$$

$$\eta \equiv e\sin(\omega), \qquad (18)$$

and the mean orbital phase,

$$\psi(t) \equiv \Omega(t - t_{asc}), \qquad (19)$$

measured from the time of ascending node $t_{asc}$, which (for small $e$) is related to $t_p$ by [51]

$$t_{asc} \equiv t_p - \frac{\omega}{\Omega}, \qquad (20)$$

and that (contrary to the time of periapse $t_p$), remains well defined even in the limit of circular orbits ($e = 0$). The parameter $\Omega \equiv \frac{2\pi}{P}$ is the mean orbital angular velocity.

The small-eccentricity phase model can be, therefore, written as

$$\phi(t) \approx 2\pi f\left\{\Delta t - a_p\left[\sin\psi + \frac{\kappa}{2}\sin 2\psi - \frac{\eta}{2}\cos 2\psi - \frac{3\eta}{2}\right]\right\}. \qquad (21)$$

## III. DATA ANALYSIS TECHNIQUE

In what follows we summarize details of the data analysis methodology used to detect CW signals hidden in the calibrated detector strain data. We commence from short FFTs of the collected data, from which we compute a time-frequency map (i.e., the peakmap). This is analyzed by applying first a thresholding criterion, and then a series of four consecutive filters with the purpose of allowing a CW signal to stand out.

### A. Short fast-Fourier-transform database

The first step consists of building up a short FFT database, the SFDB [53], where the duration of each FFT, i.e., the coherence time $T_{FFT}$, must be short enough such that the signal power remains confined within a frequency bin. The signal frequency changes in time, however, due to the Earth Doppler modulation, source spindown (if present), and -if the star is in a binary system- also due to the modulation caused by the source binary orbit. To take this modulation into account we constrain here the FFT time baseline to $T_{FFT} = 512$ s (for more details, see Appendix C). We note that the only Earth

Doppler modulation would allow us to produce FFTs longer than those used here (see, e.g., Table I of [54]).

The FFTs are then obtained from calibrated detector strain data split into interlaced (by half) chunks of $T_{\mathrm{FFT}}$ duration, each windowed in order to limit the dispersion of power due to their finite length. A time-domain data cleaning procedure, described in [53,55,56], is applied when constructing the SFDB to safely remove time-domain disturbances in detector data, which would enhance the detector noise level at the cost of a reduction in search sensitivity.

## B. Peakmap

From the SFDB we create a time-frequency map, called a peakmap, obtained selecting the most significant peaks on equalized periodograms (according to what is described in the following). This is a subtle step as the peak selection will affect the detection efficiency: all potential CW candidates, skipped at this stage due to an inaccurate construction of the peakmap, will be definitely lost. The peakmap has been described in [53,54], but here we recall only the salient aspects.

The peakmap production begins by computing, for each of the $\mathcal{N}$ FFTs in the SFDB, the ratio $\mathcal{R}$ of the square root of the periodogram [i.e., the square modulus of the $i$th FFT, $S_{P;i}(f)$] and the autoregressive average spectrum estimation, $S_{\mathrm{AR};i}(f)$ [53],

$$\mathcal{R}(i,\ell) = \sqrt{\frac{S_{P;i}(f_\ell)}{S_{\mathrm{AR};i}(f_\ell)}}; \qquad i = 1,\ldots,\mathcal{N}, \quad (22)$$

where $\mathcal{R}$ is computed for every frequency bin (indexed by $\ell$) of the $i$th FFT. By construction, the ratio $\mathcal{R}$ is an adimensional function varying around 1 and showing evident departures from 1 when spectral peaks are present.

The function $\mathcal{R}$ is compared to a threshold $\mathcal{R}_{\mathrm{th}} = \sqrt{2.5}$ [54] such that, all frequency bins above $\mathcal{R}_{\mathrm{th}}$, and that are local maxima, are selected. We call *peak* each pair consisting of a selected frequency bin and beginning time of the corresponding FFT. In general, the collection of all peaks, selected from all FFTs of the SFDB, forms the peakmap. A collection of peaks selected from a single FFT is instead referred to as a "subpeakmap."

An example of a peakmap (corrected by the Earth's Doppler modulation) is shown in Fig. 1 for simulated data covering an observation time of 30 days, and a detector power spectral density $S_h = 4 \times 10^{-24} \, \mathrm{Hz}^{-1/2}$ in the frequency band [70,73] Hz.[5] The faint tracks of three (fake) CW signals, having a gravitational-wave strain amplitude of $\sim 3 \times 10^{-24}$, and an adimensional Doppler modulation due to orbital motion of $\Delta M \sim 2 \times 10^{-4}, 6 \times 10^{-5}, 2.6 \times 10^{-4}$

---

[5]Details on how such fake data have been generated are given in Sec. VII A.
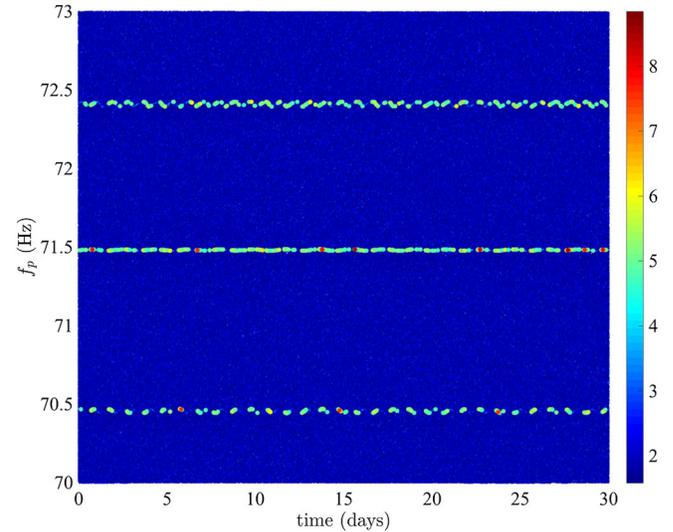


FIG. 1. Peakmap in the frequency band [70,73] Hz, where we can recognize the track of three simulated CW signals. Time is since the beginning of the run.

[see Eq. (A4) for details] for frequency of $\sim 70.5$, 71.5, and 72.5 Hz, respectively, are clearly visible as sinusoidal curves. We note that the peakmap is corrected only by the modulations caused by the Earth (orbital and rotational) motions, and not by the binary orbital motion (described by unknown parameters). If this last effect could be removed, the signals would appear as straight lines.

We stress that, contrary to what is implemented in the StackSlide [25] and Powerflux [57] schemes, we select here peaks above threshold, and that are also local maxima. This translates into a better robustness to spectral disturbances and a significant reduction of the analysis computational burden, as the number of selected peaks is smaller. This comes at the cost of only a small sensitivity loss ($\lesssim \%10$), as described in [54], where we refer the reader to for detailed statistical considerations about peak selection.

## C. Peakmap analysis

After correcting the peak frequencies for the Doppler effect due to Earth's motions, we split the frequency range into 1 Hz-wide frequency bands, and deeply analyze the peakmap in every band. Such an analysis consists of initially filtering the peak frequencies by using two standard triangular impulse response filters (with and without weighing the peak frequencies by their peak amplitude $\mathcal{R}$, respectively). Only if the maximum values of the two filter outputs are above a given threshold, the analyzed frequency bands are considered interesting, and the filtered data are convolved against a further filter, which has been built in such a way as to model the expected pattern of a CW signal.

If a signal is found, we proceed to estimate the source parameters. To this purpose, we first select the most

significant peaks,[6] and then average all the selected most significant peak frequencies corresponding to a same FFT midtime. By using those averaged peak frequencies, we can first compute a periodogram estimate and then a least-squares fit (properly described in Secs. VI A and VI B, respectively). We detail in the next section the conditions allowing us to find a signal.

## IV. IDENTIFYING SIGNAL PATTERN AND CLAIMING A DETECTION

The pattern of a CW signal, whose frequency is modulated by the source orbital motion, is characterized by two asymmetric horns [such as those hidden in the blue dots of Fig. 2(a)].[7] To identify the "double horn" feature, which is the signature of the orbital Doppler modulation, the peak frequencies in every 1 Hz analyzed frequency band undergo a set of four filters in cascade, described in the following.

Figure 2(a) shows the peak amplitude $\mathcal{R}$ as a function of the peak frequencies for two superimposed frequency bands taken as examples where a modulated CW signal is present (blue dots) and absent (red dots). In order to facilitate the task of identifying a CW signal with two asymmetric horns, we first filter the peak frequencies by using a standard triangular impulse response filter (i) [whose output is denoted with $W$ in Fig. 2(b)], and a window half-width equal to $T_{\mathrm{FFT}}^{-1}$ (empirically chosen). We then apply the same filter to the peak frequencies weighted by the peak amplitude $\mathcal{R}$ (ii), obtaining an output $W_w$ [shown in Fig. 2(c)]. We note how the double-horn feature is enhanced after applying these filters.[8]

In order to first identify frequency bands with potential signals standing out from noise, we verify if

$$\begin{cases} \max W > m(1) + 6m(2) \\ \max W_w > m_w(1) + 6m_w(2), \end{cases} \tag{23}$$

where

$$m(2) = \frac{\mathrm{median}(|W - m(1)|)}{\mathcal{C}}, \tag{24}$$

with $m(1) = \mathrm{median}(W)$, and $\mathcal{C} = 0.6745$ is a normalization factor such that, if the distribution of $W$ is normal, then $m(2)$ is the standard deviation [54]. We note that $m_w(1)$ and

$m_w(2)$ have the same meaning as $m(1)$ and $m(2)$, respectively, but are referred to $W_w$.[9]

Figure 2(b) shows the output of the triangular impulse response filter (i) applied to a 1 Hz-wide frequency band where a modulated CW signal is present (blue curve), and absent (red curve). In this last case of pure Gaussian noise, there are no outliers crossing the red straight line, which correspond to the value $m(1) + 6m(2)$.

For all frequency bands satisfying Eq. (23), we compute the ratio of $W_w/W$, illustrated in Fig. 2(d). This is done to remove the depletion visible in the blue curves of Figs. 2(b) and 2(c), which is due to the peak selection effect in the peakmap [53,54], and is more prominent for loud signals. To enhance the signal contribution, the ratio $W_w/W$ is then convolved against a filter (iii), which we have *ad hoc* built with a nonsymmetrical Gaussian shape response, and finally a filter identical (iv), but running in opposite direction. This is needed to reproduce, in the plane $(f, \mathcal{R})$, the shape of a signal whose frequency is modulated by the source orbital motion [see blue dots in Fig. 2(a)].

Such a Gaussian-like filter, modeling half of the horn-shaped signal modulation pattern, is given by

$$\begin{cases} \mathcal{G}(1:\mu) = e^{-\frac{(t(1:\mu)-\mu)^z}{u\sigma^2}} \\ \mathcal{G}(\mu+1:n_t) = e^{\frac{(t(\mu+1:n_t)-\mu)^q}{g\sigma^2}}, \end{cases} \tag{25}$$

where $\mu = 3\sigma$, $n_t = 12\sigma$, $z = 2$, $u = 2$, $q = 1.7$, $g = 10$, and the "standard deviation" $\sigma = 6$ (all values found empirically).

We refer the reader to Appendix E for further details on the Gaussian-like filter.

To model the remaining half part, the convolution is computed by using the same Gaussian-like filter running in opposite direction.

In general, if a signal is present in a band, when the Gaussian-like filter runs in one direction, only one of the horns in Fig. 2 gets amplified, and the other horn will be amplified by the filter running in opposite direction. If this condition is not satisfied, we consider the signal *not found* in the analyzed frequency band. Alternatively, we can *claim a detection* and identify the band around the modulated signal pattern. For safety reasons we select, however, a wider band consisting of roughly 100 additional samples (corresponding to ∼10 mHz) on both sides.

The standard deviation of the Gaussian-like filter $\sigma = 6$ is chosen to be wide enough to account for the Doppler shift due to source orbital motion. If a signal is not found in a frequency band, however, a second attempt is done by convolving the data against an identical Gaussian-like filter,

---

[6] We call most significant peaks those having a peak amplitude $\mathcal{R}$ larger than the threshold value of ∼3.5 (see Appendix D for details).

[7] We note that such an asymmetry depends on the combination of orbital parameters, and is also present for circular orbits.

[8] The filters (i) and (ii) are implemented in the SNAG software package [58], with the possibility of choosing either a triangular, rectangular or exponential window, which exhibit similar results when applied to the same data set. We opted, however, for the most intuitive choice of a triangular impulse response.

[9] We use the median, rather than the mean, as it is more robust in the presence of outliers, and we choose it to build a robust estimator of the dispersion parameter, which we use instead of the classical standard deviation.

(a)

(b)

(c)

(d)
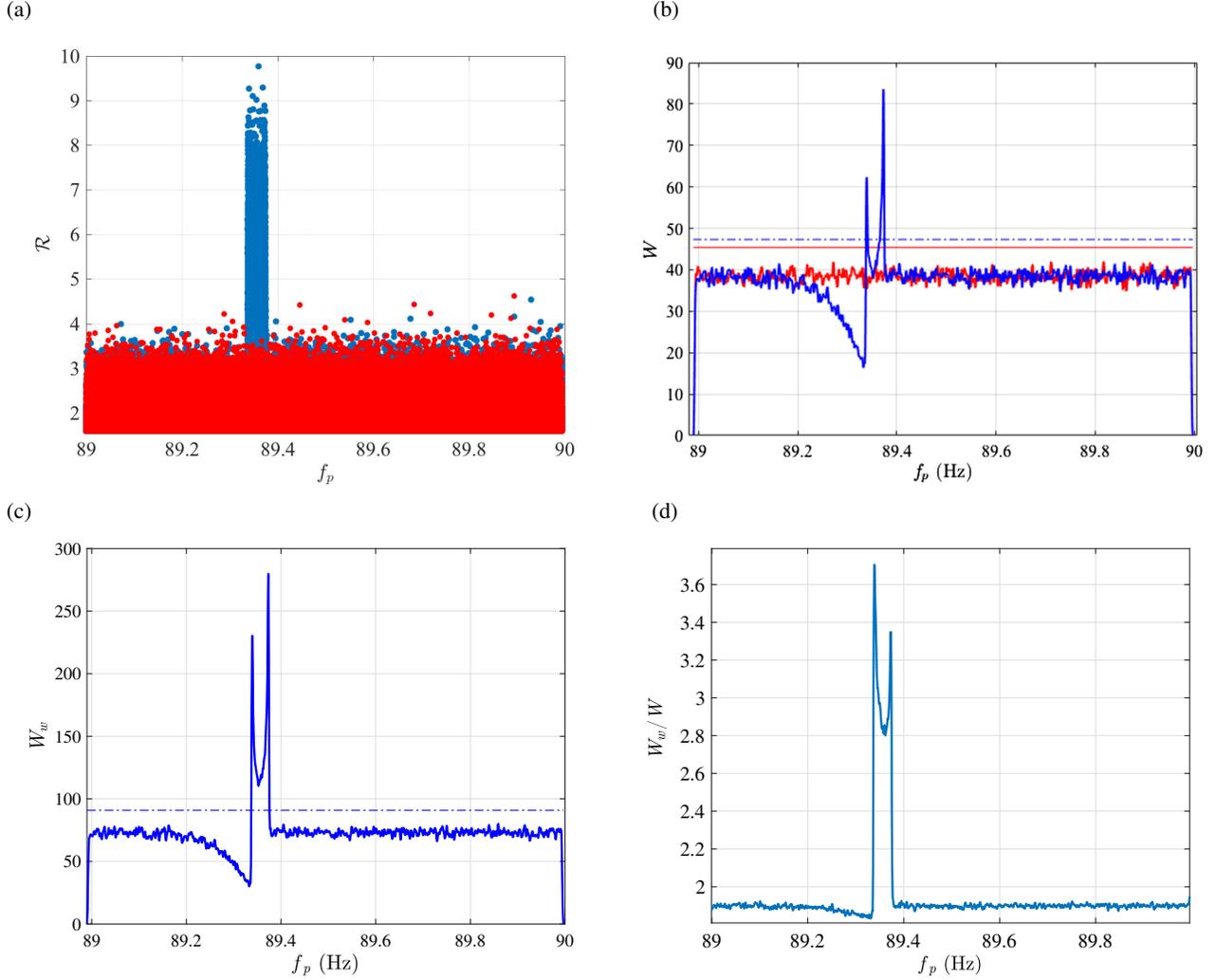
FIG. 2.    Panel (a): Peak amplitude $\mathcal{R}$ versus peak frequencies for two frequency bands where a modulated CW signal is present (blue dots), with an orbital Doppler modulation $2\Delta M f_s \sim 37$ mHz (see Appendix A for details), and absent (red dots). Panel (b): Output $W$ of a triangular impulse response filter applied to the peak frequencies in a band where there is a signal added to Gaussian noise data (blue curve), and the same band when no signal is present, but just pure Gaussian noise (red curve). The highest and lowest lines correspond to the value $m(1) + 6m(2)$ in the two cases, respectively. Panel (c): Output $W_w$ of the triangular impulse response filter applied to the peak frequencies weighted by their respective peak amplitude; the dashed line corresponds to the value $m_w(1) + 6m_w(2)$ (see main text for the definition of the $m$ and $m_w$ values.). Panel (d): Ratio of $W_w/W$ versus peak frequencies. We note that the signal visible in Fig. 2(a) has a strain amplitude of $\sim 2.7 \times 10^{-24}$ over a $T_{\mathrm{obs}} = 30$ days. As shown in Sec. X, we would have detected the same signal also if it would have had a strain amplitude of $\sim 8.7 \times 10^{-25}$ over the same observation time.

but with $\sigma = 1$, to take into account also smaller Doppler modulation effects.

We stress that what we refer to as detection claim here corresponds to candidates (i.e., potential signals) that must however undergo deep follow-up studies, such as [30,59], before being eventually confirmed as real detections.

## V. DETAILS OF THE SEARCH METHOD

The new strategy presented here relies on the basic consideration that the frequency modulation caused by the source orbital motion can be used to unveil the signature of CW signals, and to also extract information on source

parameters. We summarize the salient steps of the procedure in the following:

(1) We start from the subpeakmaps obtained by processing the artificial data set generated according to the guidelines described in Sec. VII. In particular, we have $\mathcal{N} = 10127$ subpeakmaps and interlaced FFTs of duration $T_{\mathrm{FFT}} = 512$ s, covering an overall one month of gapless data.

*Peakmap analysis:*

(2) For every subpeakmap, for the source sky location assumed to be known with enough accuracy, we apply a frequency correction to account for the Doppler shift due to the Earth motions. Hence,

neglecting relativistic effects, we shift the peak frequencies $f_p$ from the original received ones, $f_p^o$, according to

$$f_p = \frac{f_p^o}{(1 + \frac{\vec{v}\cdot\vec{n}}{c})}, \qquad (26)$$

where $\vec{n}$ is the unit vector pointing from the SSB to the source (see Sec. II); $\vec{v}$ is the detector velocity with respect to the SSB frame, and is given by the sum of two components, from the yearly Earth motion around the Sun and from the rotation of Earth around its axis. These velocities are computed at the FFT midtimes ($t_c$) by using the DE405 JPL Solar System Ephemeris [60,61].

We emphasize that $f_p$ are then the peak frequencies modulated due to the only source orbital motion, which is the information we need for the current procedure.

(3) We analyze the search frequency range splitting it into 1 Hz frequency bands, and each of them undergoes scrutiny to establish if the modulated pattern of a CW signal is present. To this purpose we apply the filters described in Sec. IV and, if a signal is found, we extract the subband containing the modulated pattern that needs additional inspection for the parameter estimation.

(4) In every subband identified as above, we select the most significant peaks crossing the threshold established in Appendix D (i.e., having $\mathcal{R} > \theta_{\text{thr}} \sim 3.5$).

(5) We then average the frequencies of the most significant peaks identified at the previous stage, and corresponding to a same FFT midtime, resulting at most in $\mathcal{N}$ ($t_c, \bar{f}_p$) pairs per subband.[10]

*Parameter estimation:*

(6) For each subband, we perform a periodogram estimate (detailed in Sec. VI A) for the unevenly spaced data set ($t_c, \bar{f}_p$) in order to look for periodicities, and possibly estimate source orbital period.

(7) We then carry on with a least-squares fitting of sine waves to estimate the signal frequency and remaining orbital parameters (as depicted in Sec. VI B).

Further details follow to describe steps 6 and 7.

---

[10]In the current analysis we found a minimum (maximum) of 14 (9 417) pairs ($t_c, \bar{f}_p$), in the band [151,152] Hz ([82,83] Hz), where a signal with an amplitude $h_0 = 1.1 \times 10^{-24}$ ($h_0 = 4.3 \times 10^{-24}$) has been injected. Special care must be devoted not to consider such a number of FFTs as those contributing to detect a putative signal, but rather to estimate the parameters of a previously identified signal. In fact, a detection claiming (defined in Sec. IV) depends only on the performance of the filters introduced in Sec. IV, and applied before selecting peaks with $\mathcal{R} > \theta_{\text{thr}}$.

## VI. PARAMETER ESTIMATION METHODOLOGY

### A. Spectral analysis of unequally spaced data

If a signal is found, we select the frequency band around the modulated signal pattern, and proceed on selecting only the most significant peaks[11] within such a subband. Then we average all the selected most significant peak frequencies corresponding to a same FFT midtime.

For all subbands where a putative signal has been identified, we face the problem to sift through the unevenly sampled data ($t_c, \bar{f}_p$) in order to unveil any potential periodicity. In general, Fourier analysis is employed to characterize the frequency content of a signal, thus detecting possible periodicities. The FFT algorithm computes a Fourier decomposition under the assumption that the input data points are equally spaced in the time domain. However, approximate Fourier transform methods have also appeared in the literature, which treat nonequispaced data (see, e.g., [62]). More generally, there are other methods used to perform a spectral estimate of nonregularly spaced data, which are based on periodogram analysis.

The periodogram estimate is well suited to the problem of detecting a periodic signal in the presence of noise, and is the most common method to estimate the power spectrum of evenly and unevenly spaced data, due to the simplicity of its statistical behavior [see Eqs. (4) and (10) of [63] for evenly and unevenly spaced data, respectively].

### 1. Peak power spectrum estimate and orbital period recovery

In the following we illustrate an alternative, and more straightforward, estimate of the power spectrum[12] with respect to the standard Lomb-Scargle periodogram estimator. We refer the reader to [63,64] for more details, including the statistical properties of the Lomb-Scargle periodogram.

For every frequency subband (selected within every 1 Hz analyzed band), where a signal has been pinpointed, we estimate the peak power spectrum as follows,

$$\mathcal{S}(\nu_j) = \frac{1}{N} \left| \sum_{k=1}^{N} Y_k e^{2\pi i \nu_j t_{c,k}} \right|^2, \qquad (27)$$

---

[11]As already said, the most significant peaks are those with $\mathcal{R} > \theta_{\text{thr}} \sim 3.5$.

[12]We note that we use here "power spectrum" and "periodogram" synonymously, although specifically the power spectrum is a theoretical quantity defined as an integral over continuous time, of which the periodogram is purely an estimate based on a amount of discrete data.

where $Y_k = \bar{f}_{p,k} - \langle \bar{f}_p \rangle_N$ are the deviations of $\bar{f}_{p,k}$ from their mean value,[13] and $\bar{f}_{p,k}$ are the $k$th averaged peak frequencies. The number $N$ of averaged frequencies varies for every frequency subband and is at most $N = \mathcal{N}$ when the signal contribution comes from all subpeakmaps. We recall that the times $t_{c,k}$ correspond to the midtimes of the $k$th FFT (or, equivalently, subpeakmap) in the analyzed frequency subband.

The number $N_\nu$ of frequency points $\nu_j$ (with $j = 1, \ldots, N_\nu$) at which to compute the periodogram is given by

$$N_\nu = \frac{\nu_{\max} - \nu_{\min}}{d\nu}, \qquad (28)$$

where $\nu_{\min} = 0 \text{ d}^{-1}$, $\nu_{\max} = 10 \text{ d}^{-1}$, and the step $d\nu = 1/(T_{\text{obs}} r_\nu)$, with $r_\nu$ being a refinement factor. The values of $0 \text{ d}^{-1}$ and $10 \text{ d}^{-1}$ correspond to considering orbital periods from a few hours ($\sim 2.4$ h) up to infinity. This very wide range includes of course the source orbital periods we target, which range from 10 h up to 48 h.

The resolution at which to evaluate the periodogram can be either refined or coarsened via $r_\nu$. A reasonable choice is a refinement in frequency resolution $r_\nu = 4$. Hence, for the observation time $T_{\text{obs}} = 1$ month considered here, we compute the periodogram for 1200 frequencies.

By taking the maximum value of $\mathcal{S}(\nu_j)$, and the inverse of the fundamental frequency, which is the frequency corresponding to max $\mathcal{S}(\nu_j)$, i.e., $\nu^m$, we obtain the source orbital period, namely

$$P = \frac{1}{\nu^m}. \qquad (29)$$

Appendix F illustrates an example of the periodogram appearance, and how the orbital eccentricity impacts on the number of harmonics. Perfect consistency with the Lomb-Scargle periodogram is also discussed.

### B. Least-squares fitting of sine waves

We illustrate here the method employed to estimate the signal frequency and all source orbital parameters but the orbital period.

For every subband of interest, we perform a sinusoidal fit of the averaged peak frequencies $\bar{f}_p$,

$$\bar{f}_{p,k} = A_0 + \sum_{h=1}^{N_h=2} [A_{2h-1} \cos(\Omega h t_{c,k}) + A_{2h} \sin(\Omega h t_{c,k})], \qquad (30)$$

---

[13] The mean is a marginal statistic and is not important for judging periodicity. Hence, we can safely subtract the sample mean in Eq. (27) in order to obtain a series with zero mean. Not subtracting the sample mean would make the scale of the periodogram plot difficult to judge, mainly if $\mathcal{S}(0)$ is very large [65].

with $k = 1, \ldots, N$, $\Omega = 2\pi/P$, and $P$ being the source orbital period recovered by the periodogram technique detailed in Sec. VI A 1. The $A$ coefficients are real numbers with dimension of the inverse of a time.

In Appendix G, we provide details on the spectral content of the source orbital modulation, i.e., $\dot{R}(t)$, and we note (from Fig. 15) that only two harmonics are necessary to describe the spectral content of $\dot{R}/c$. Hence, we chose $N_h = 2$.

We then have to solve the overdetermined linear system of $N > N_h$ equations, given by Eq. (30), in the unknowns $A_{0,1,2,3,4}$, i.e.,

$$\mathcal{A} \cdot X = \mathcal{Y}, \qquad (31)$$

with

$$\mathcal{A} = \begin{bmatrix} 1 & \cos(\Omega t_{c,1}) & \sin(\Omega t_{c,1}) & \cos(2\Omega t_{c,1}) & \sin(2\Omega t_{c,1}) \\ 1 & \cos(\Omega t_{c,2}) & \sin(\Omega t_{c,2}) & \cos(2\Omega t_{c,2}) & \sin(2\Omega t_{c,2}) \\ . & . & . & . & . \\ . & . & . & . & . \\ . & . & . & . & . \\ 1 & \cos(\Omega t_{c,N}) & \sin(\Omega t_{c,N}) & \cos(2\Omega t_{c,N}) & \sin(2\Omega t_{c,N}) \end{bmatrix}, \qquad (32)$$

$$X = \begin{bmatrix} A_0 \\ A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix}, \qquad \mathcal{Y} = \begin{bmatrix} \bar{f}_{p,1} \\ \bar{f}_{p,2} \\ . \\ . \\ . \\ \bar{f}_{p,N} \end{bmatrix}. \qquad (33)$$

The exact solution is obtained by using the least-squares method.

Equation (30) is an equality between $\bar{f}_{p,k}$ and a sum of sinusoids at frequencies $1/P, 2/P$, with $1/P$ the first harmonic (i.e., the fundamental frequency). Technical details on how establishing the existence of these two peak power spectrum harmonics can be found in Appendix H. The amplitudes of the first and second sinusoid in Eq. (30) are given by $\mathcal{H}_1 = \sqrt{A_1^2 + A_2^2}$ and $\mathcal{H}_2 = \sqrt{A_3^2 + A_4^2}$, respectively, while $A_0$ is the so-called DC value. This term corresponds to the amplitude of a cosine wave with zero frequency [note, in fact, the presence of ones in Eq. (32)].

In order to solve Eq. (31), we multiply each member of such an equation by $\mathcal{A}^T$, with the superscript $T$ denoting the transpose matrix, and obtain

$$\mathcal{B} \cdot X = \mathcal{D}, \qquad (34)$$

where $\mathcal{D} = \mathcal{A}^{\mathrm{T}} \cdot \mathcal{Y}$, and $\mathcal{B} = \mathcal{A}^{\mathrm{T}} \cdot \mathcal{A}$ is a $(5 \times 5)$ square matrix. The number 5 comes from considering the $b = 2N_h + 1$ amplitudes $A_{0,1,2,3,4}$ (chosen $N_h = 2$).

The estimate of the unknown $X$ parameters is obtained by solving Eq. (34), i.e.,

$$X = \mathcal{B}^{-1} \cdot \mathcal{D}; \qquad (35)$$

$X$ is the least-squares solution that minimizes the $(N - 5)$ degrees-of-freedom $\chi^2$ variable

$$\chi^2 = \sum_{k=1}^{N} (\bar{f}_{p,k} - \mathcal{A} \cdot X)^2, \qquad (36)$$

where $(\bar{f}_{p,k} - \mathcal{A} \cdot X)$ are the residuals, whose mean indicates the accuracy of the solution found: the closer to 0, the more accurate $X$. The inverse matrix $\mathcal{C} = \mathcal{B}^{-1}$ is the covariance matrix of the $A_{0,1,2,3,4}$ parameters, and the elements on the diagonal of $\mathcal{C}$ are proportional to the variance of $A_{0,1,2,3,4}$. Hence, their uncertainties are

$$dA_b = \sqrt{\mathcal{C}_{bb}} \frac{1}{T_{\mathrm{FFT}}}, \quad b = 0, \ldots, 4. \qquad (37)$$

Having then solved Eq. (35), we can estimate the $A_{0,1,2,3,4}$ parameters and explicitly write Eq. (30) for every $k$ as

$$\bar{f}_{p,k} = \boxed{A_0} + [\, \boxed{A_1 \cos(\Omega\, t_{c,k}) + A_2 \sin(\Omega\, t_{c,k})} + \boxed{A_3 \cos(2\Omega\, t_{c,k}) + A_4 \sin(2\Omega\, t_{c,k})}\,], \qquad (38)$$

which we can compare against the received frequencies, which are modulated due to the source orbital motion, i.e.:

$$f_m = f - f \frac{\dot{R}(t_{c,k})}{c} = \boxed{f} - f a_p \Omega [\, \boxed{\cos(\Omega(t_{c,k} - t_{\mathrm{asc}}))} + \boxed{\kappa \cos(2\Omega(t_{c,k} - t_{\mathrm{asc}})) + \eta \sin(2\Omega(t_{c,k} - t_{\mathrm{asc}}))}\,], \qquad (39)$$

with the orbital Doppler modulation contribution obtained by deriving Eq. (16) with respect to time.

From the comparison of the dark-gray highlighted terms of Eqs. (38) and (39), we find the signal frequency

$$f = A_0. \qquad (40)$$

By performing a fit with a sinusoidal function, i.e., by comparing the gray highlighted terms of Eqs. (38) and (39), we obtain an estimate of the projected orbital semimajor axis:

$$a_p = \frac{\mathcal{H}_1}{A_0 \Omega}. \qquad (41)$$

Finally, by comparing the light-gray highlighted terms, we find the orbital eccentricity given by the ratio of the two harmonic amplitudes:

$$e = \frac{\mathcal{H}_2}{\mathcal{H}_1}. \qquad (42)$$

The argument of periapse is given by

$$\omega = \arctan\left(\frac{A_4}{A_3}\right), \qquad (43)$$

where we used that $e = \sqrt{\kappa^2 + \eta^2}$ and $\omega = \arctan(\eta/\kappa)$, due to Eqs. (17) and (18).

Finally, we can compute the time of periapse from Eq. (20):

$$t_p = \frac{1}{\Omega}\left[\omega - \arctan\left(\frac{A_2}{A_1}\right)\right], \qquad (44)$$

being

$$t_{\mathrm{asc}} = -\frac{1}{\Omega} \arctan\left(\frac{A_2}{A_1}\right). \qquad (45)$$

## VII. SIGNAL GENERATION

We describe the steps performed to create a fake data set to which we added simulated CW signals, which have the scope to validate the new presented algorithm. The choice of reasonable ranges from which the source parameters are drawn is also discussed.

### A. Data production

We have generated one month of gapless Gaussian detector noise data assuming the expected best strain sensitivity of advanced LIGO-Virgo [11–13] detectors, i.e., a noise spectral density $S_h = 4 \times 10^{-24}$ Hz$^{-1/2}$ approximately constant in the frequency interval [70, 200] Hz. We have produced $\mathcal{N} = 10127$ interlaced FFTs with duration $T_{\mathrm{FFT}} = 512$ s each.[14] Then, we artificially generated and injected into (i.e., added to) such a data set 131 CW signals emitted from pulsars in low- and high-eccentricity binary systems. Software injections have been performed using the LALSUITE software package [66]. As already anticipated, we have neglected spin wandering

---

[14]The FFT duration is in general obtained by requiring the signal power to be concentrated in less than a frequency bin (see Appendix C for more details). Furthermore, we specify that the FFT production is a common piece of other CW analyses.

effects since the current method—due to its robustness—has a high tolerance to small frequency variations and then is expected not to be limited by possible spin wandering.

## B. Choice of signal parameters

In Fig. 3, we plot the distribution of orbital eccentricities versus the projected semimajor axis for 221 ATNF catalog pulsars found in binary systems. The minimum and maximum orbital periods in Fig. 3 are roughly 1.6 h and 46 y, respectively. The binary systems we are interested in, however, have orbital periods $10 \text{ h} \leq P \leq 2 \text{ d}$, projected semimajor axis $a_p$ of a few seconds, and no restriction on orbital eccentricity (see magenta stars in Fig. 3).[15]

The so-called signal amplitude parameters are randomly chosen from uniform distributions as follows: the scalar gravitational-wave amplitude $h_0^s \in [1, 5] \times 10^{-24}$, the inclination angle $\cos \iota_s \in [-1, 1]$, the polarization angle within $\psi_s \in [0, 2\pi]$, and the (irrelevant) initial phase within $\phi_{0_s} \in [0, 2\pi]$.[16]

The sky-position for all signals is fixed to that of Scorpius X-1, namely $(\alpha_s, \delta_s) = (4.276, -0.273)$ rad.[17]

The so-called signal phase-evolution parameters $\{f_s, a_{p_s}, t_{p_s}, P_s, e_s, \omega_s\}$ are generated by randomly drawing them from uniform distributions over the ranges:

$$f_s \in [70, 200] \text{ Hz}$$
$$a_{p_s} \equiv \frac{a \sin i}{c} \in [1, 3] \text{ s},$$
$$P_s \in [10, 48] \text{ h},$$
$$t_{p_s} \in \left[ t_{\text{mid}} - \frac{P}{2}, t_{\text{mid}} + \frac{P}{2} \right],$$
$$\log_{10} e_s \in [-6, \log_{10}(0.9)],$$
$$\omega_s \in [0, 2\pi] \text{ rad}, \qquad (46)$$

with $t_{\text{mid}}$ being the midtime of the whole observation.

We remark that we consider also high-eccentricity orbits, favoring low-eccentricity ones as more copious (see Fig. 3).

Every frequency $f_s$ has been actually uniformly drawn from a subinterval of 0.4 Hz around the mid interval of every analyzed 1 Hz band (which is larger than the modulation expected from the orbital motion of the sources considered here). This is done to avoid border effects.[18]

---

[15]We note that such a class of signals includes, of course, Scorpius X-1.

[16]The initial phase $\phi_0$ at the reference time $t_{\text{ref}}$ has to be added in Eqs. (13) and (21), but for our purposes it can also be neglected.

[17]In order to avoid sensitivity losses from sky positions which are less favorable at certain times, the selected peaks can be normalized with *ad hoc* weights based on the so-called antenna pattern, i.e., the directional (angular) response of the detectors. When necessary, this can be done in a way analogous to what is implemented in [54].

[18]In the presence of signals whose frequency modulation spans more than one frequency bin, the SNR would be lessened in every bin.
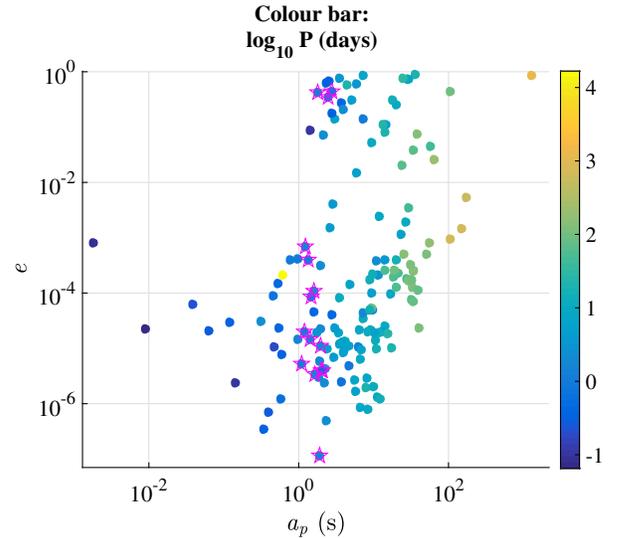


**Colour bar:**
$\log_{10} \text{P (days)}$

FIG. 3. Distribution of orbital eccentricities versus projected semimajor axis for 221 ATNF catalog pulsars found in binary systems with rotation frequencies $\geq 10$ Hz, and up to $\sim 716$ Hz [18,67]. The colour bar indicates the logarithmic values of source orbital periods. The points surrounded by magenta stars correspond to sources with $10 \text{ h} \leq P \leq 2 \text{ d}$, $1s \leq a_p \leq 3 \text{ s}$, and $0 < e < 1$.

which would make borderline detections more difficult. A way to circumvent this issue consists of performing a trivial interlacement of all analyzed frequency intervals, which will be anyway done in a real search. This would increase the computing cost, as it would double the number of analyzed frequency bands, but will avoid to underestimate signal amplitudes and miss borderline detections. Such an increase in computing time will anyway keep the method quite computationally feasible even on a single-processor computer.

In Fig. 4, we show the orbital eccentricities versus frequency for 131 simulated sources (light dots) and 45 known pulsars found in binary systems (dark asterisks) [18,67] in the same frequency and eccentricity ranges of the simulated population. The asterisks surrounded by circles indicate sources with $10 \text{ h} \leq P \leq 2 \text{ d}$ and $1 \leq a_p \leq 3 \text{ s}$, which is our target population. Although they are only three sources, there are several other pulsars we expect to exist in the Galaxy, which have not yet been discovered (as discussed in Sec. I). Hence, it becomes crucial to sift a parameter space as much large as possible, without exclusion of particular regions.

As stated earlier, we split the searched frequency range into smaller frequency bands that will be deeply analyzed to find evidence of CW signals. We choose 1 Hz as width to analyze the various frequency bands, which is much larger than the maximum modulation caused by the orbital motion of the source in the parameter space investigated here. In fact, Fig. 5 exhibits the maximal Doppler shift due to orbital motion, i.e., $2f_s \Delta M$ (discussed in Appendix A), versus frequency of 131 simulated sources. We see that $2f_s \Delta M$ varies from $8.2 \times 10^{-3}$ Hz up to $\sim 0.2$ Hz.
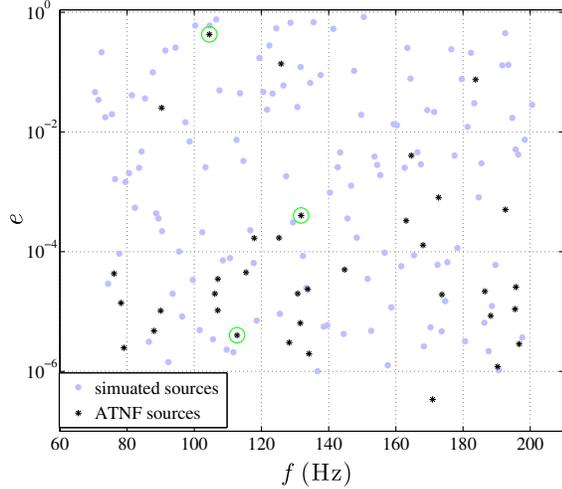
FIG. 4. Distribution of orbital eccentricities versus frequency for the population of 131 simulated sources (dots) and 45 ATNF catalog pulsars found in binary systems (asterisks). The asterisks surrounded by empty circles correspond to sources with orbital periods and projected semimajor axes compatible with our target population.
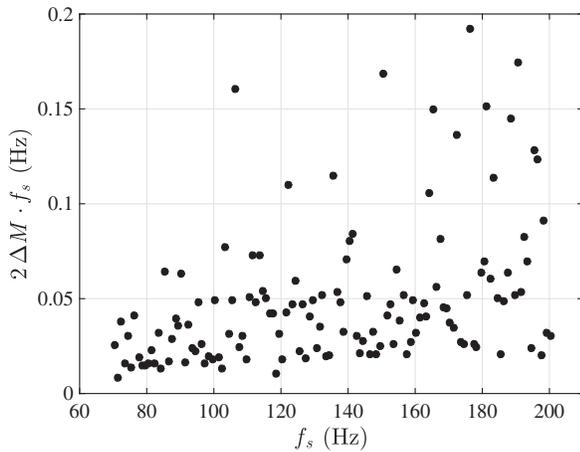


FIG. 5. Maximal Doppler shift due to orbital motion ($2f_s \Delta M$) as a function of the frequency for the population of 131 simulated sources.

## VIII. RESULTS: DETECTION AND PARAMETER ESTIMATION

Based on the condition outlined in Sec. IV, we can claim 128 signal detections out of total 131. The frequency bands where our algorithm fails to detect a signal are [141,142] Hz, [166,167] Hz, and [190,191] Hz. The gravitational wave strain amplitudes of the three not-detected signals is $h_0 \sim 10^{-24}$ for one month of single-detector data set we process. The signals are missed as they are not loud enough to stand out of the background noise, mainly due to a large orbital Doppler modulation.

Figure 6 shows the parameter estimates for 128 detected signals; the parameters of the three signals that could not be

detected are denoted by filled diamonds in all panels. The offsets on the vertical axes give the absolute value of the difference between estimated and true signal parameters (denoted with $r$ and $s$ subscripts, respectively),

$$\Delta \mathcal{O} = \mathcal{O}_r - \mathcal{O}_s, \tag{47}$$

where $\mathcal{O}$ refers to a generic parameter among those shown in Fig. 6.

We observe that the recovery of signal frequency $f_s$, projected orbital semimajor axis $a_{p_s}$, and orbital period $P_s$ is generally quite good. The behavior of the one-sigma orbital period uncertainty $dP$ [inferable by the open squares in panel (c) of Fig. 6] is larger than typical actual offsets and reflects the scaling of $dP$ with $P^2$, i.e.,

$$dP = P^2 d\nu, \tag{48}$$

with $d\nu = (4T_{\text{obs}})^{-1}$, as described in Sec. VI A 1. Obviously, the higher the orbital period, the less precise its estimate as the number of orbits observed during $T_{\text{obs}}$ decreases. The offsets larger than $10^4$ s in Fig. 6(c) correspond to three signals where the periodogram estimate does return an orbital period which is half of the true value. This happens for sources with an orbital period close to multiples and submultiples of the Earth's sidereal day, especially if they are not strong enough to stand out of noise and have a quite high eccentricity. In these circumstances, multiples and submultiples of the Earth periodicity can be mistaken for the signal periodicity. A way to bypass this contingency is to perform a sinusoidal fit with two harmonics (similarly to what is discussed in Sec. VI B), looping over several target orbital periods and then choosing the orbital period value for which the fit exhibits the smallest residuals. Although such approach will increase the computational cost (e.g., if a large number of orbital periods is considered), the method will remain computationally affordable (on a single-processor computer) and will improve even further the source parameter estimation. The investigations of this technique will be performed in a separate study.

The uncertainties on the source parameters are obtained by using standard error propagation rules for uncorrelated variables from Eqs. (40), (41), (42), (43), (44) and using as $dA_b$ uncertainties (with $b = 0, \ldots, 4$) those given by Eq. (37).

The estimate of signal eccentricity $e_s$, argument of periapse $\omega_s$, and time of periapse passage $t_{p_s}$ is more delicate. The crosses in panels (d) and (e) of Fig. 6 correspond to 69 cases where a signal is detected but no estimate can be provided for $e_s$ and $\omega_s$, as the second harmonic is not found in the corresponding periodogram, and hence the recovered eccentricity would be indistinguishable from zero (as also described in Appendix G). If
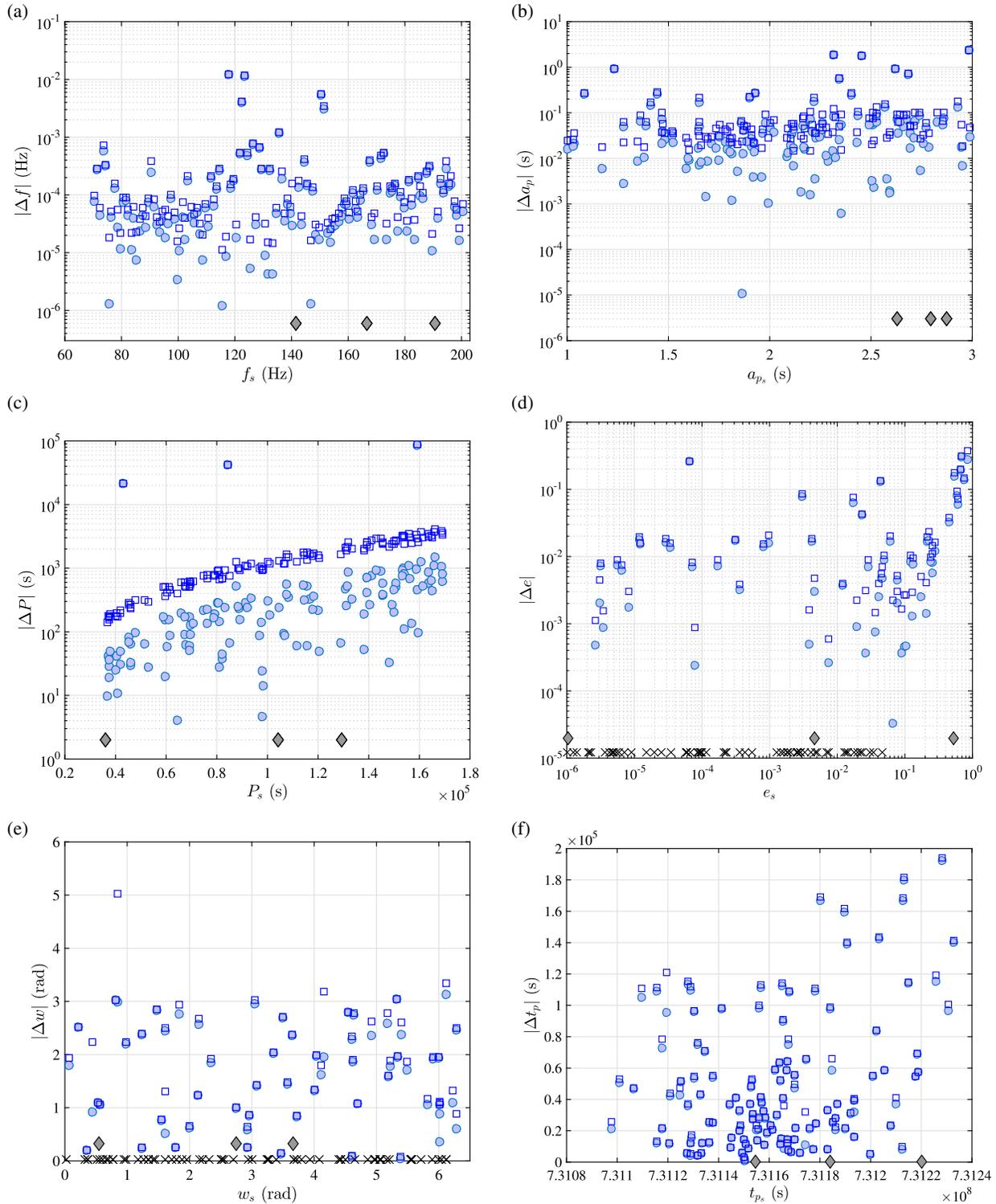
FIG. 6. Parameter estimation for detected signals as a function of 131 injected signal (a) frequency $f_s$, (b) projected orbital semimajor axis $a_{p_s}$, (c) orbital period $P_s$, (d) eccentricity $e_s$, (e) argument of periapse $\omega_s$, and (f) time of periapse passage $t_{p_s}$. The light-blue filled circles show the absolute value of the difference between the injected (true) and recovered values of signal parameters. The empty squares show the sum of the offsets $|\Delta\mathcal{O}|$ and the quoted one-sigma errorbars corresponding to each estimate. The filled diamonds mark the parameters of signals that have not been detected (and for which no parameter estimates can be provided). The crosses in panels (d) and (e) indicate the parameters of 69 signals for which no estimate can be obtained as the second harmonic is not found in the corresponding periodogram. Hence, the estimate is provided only for 59 signals.

the second harmonic does not exist,[19] the current method does not have the ability to estimate the eccentricity. Further studies are needed to determine the best strategy to use for such situations.[20]

The $t_{p_s}$ estimate is generally poor and would improve with more precise estimates of $\omega_s$ and $P_s$.

We highlight that in Eq. (39) we use the small-eccentricity approximation for $\dot{R}/c$, which also works well overall to estimate high eccentricities. Future studies, however, will be devoted to understanding how to relax this condition (by making suitable use of Eq. (G1) in a least-squares fitting approach).

The novel method presented here allows us to detect the bulk (~97%) of injected signals and to perform satisfactory parameter estimates by using a standard portable computer. All source parameter estimates, however, can be further improved by using a hierarchical approach, which use the current parameter estimates to obtain an approximate demodulation of the source orbital motion, and then produce longer FFTs, by assuming such a partially known modulation. Then, the same procedure can be reiterated on the new data set, thus improving the parameter estimation and search sensitivity by a factor depending on the square root of how much the FFT duration can be increased. Such approach, however, will be implemented in a separate pipeline.

This is the first algorithm in the literature able to provide estimates for orbital eccentricity and the argument of periapse.

## IX. RECOVERY OF SIGNAL STRAIN AMPLITUDE

The current search is sensitive to a combination of $h_0$ and $\cos\iota$, which is given by [23]

$$H_0 = \frac{h_0}{2}\sqrt{1 + 6\cos^2\iota + \cos^4\iota}. \tag{49}$$

We can define a theoretical SNR in terms of the $H_0$ amplitude,

---

[19]As a reminder, we used the criterion discussed in Appendix H to check if the second harmonic of the periodogram exists.

[20]Attempts to discard the peaks due to noise will be investigated. Once the first harmonic is well characterized in terms of amplitude, frequency and phase (i.e., a sinusoid), a viable approach might consist of determining a range of possible values around it, i.e., a minimum and maximum sinusoid around the first harmonic in the peak frequency and time plane. These "boundary sinusoids" can be obtained by the periodogram background and can be used to discard all the peaks that do not lie in such a region. At this point, the surviving peak-frequencies can be averaged and a new periodogram can be computed. This new periodogram will have a reduced noise background with respect to the original one, and the second harmonic should more likely stand out of noise. Detailed studies, which we postpone to future work, must be however performed to characterize the boundary sinusoids and accurately test this approach.
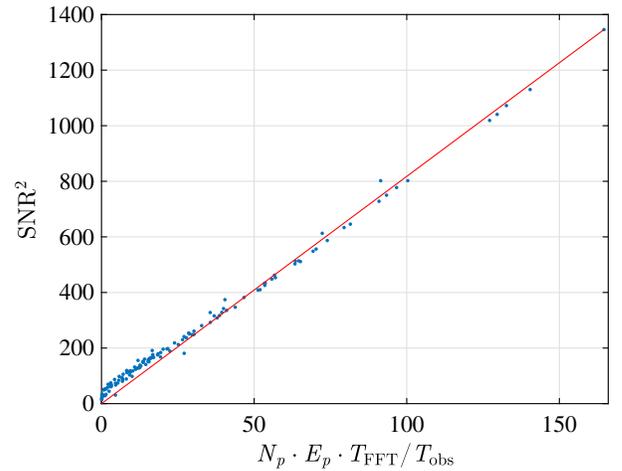


FIG. 7. Observed $\mathrm{SNR}^2$ versus $N_p E_p T_{\mathrm{FFT}}/T_{\mathrm{obs}}$ (points), and related linear fit (straight line).

$$\mathrm{SNR} = H_0\sqrt{\frac{T_{\mathrm{FFT}}}{S_h}}, \tag{50}$$

and find an empirical relation between such SNR and the number of peaks selected above threshold, on average, per FFT in a frequency band where a signal has been identified that are properly weighed by taking into account the signal amplitude, i.e.,

$$\mathrm{SNR}^2 = N_p E_p \frac{T_{\mathrm{FFT}}}{T_{\mathrm{obs}}} + \mathrm{corr}. \tag{51}$$

We labeled $N_p$ as the number of peaks in a small frequency band where a CW signal has been found, and being above the threshold $\mathcal{R} > \theta_{\mathrm{thr}} \sim 3.5$ (see Appendix D for more details), while $E_p$ is the average of the squared peak amplitude above the same threshold in that band.[21]

Figure 7 shows the relation expressed by Eq. (51) [with the SNR given by Eq. (50)], which is linear apart from a correction factor "corr" that is necessary to account for the discrepancies obtained for low SNR signals. We note that this correction factor is included in the uncertainty estimates provided in Fig. 8.

In the top (bottom) panel of Fig. 8 we show the consistency of the estimated signal amplitudes $H_0^r$ ($h_0^r$) with the true values $H_0^s$ ($h_0^s$), together with the relative uncertainties that are especially small.

We emphasize that such a method can be used to estimate the gravitational-wave strain amplitude also in CW searches for isolated NSs, as there is no dependency on orbital parameters. Hence, by using Eqs. (50) and (51), and interpolation (or extrapolation, if needed) schemes, for a given set of $N_p$, $E_p$, $T_{\mathrm{FFT}}$ and $T_{\mathrm{obs}}$, we can find the $H_0$

---

[21]The frequency band of the modulated signal pattern is identified according to what is outlined in Sec. IV.
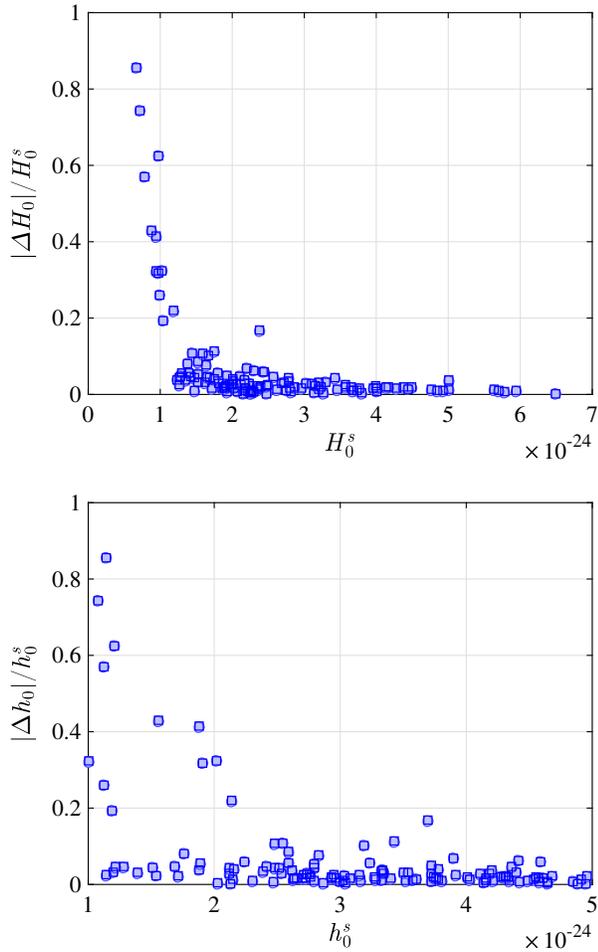
FIG. 8.  Fractional errors in strain amplitude estimates divided by the true values (blue filled circles) and, almost superimposed, the same errors summed to the one-sigma error bars corresponding to each estimate (empty squares). The estimates are provided for 128 out of 131 original inspected frequency bands, where an artificial CW signal has been injected and detected (in each of them).

corresponding to a given signal, which has been previously identified in a frequency band.

## X. SENSITIVITY ESTIMATE

The customary *modus operandi* to rigorously estimate the sensitivity of a search for CW signals is based on cumbersome Monte-Carlo simulations, which require consecutive signal injections. We circumvent such expensive approach, opting for a less accurate, but expeditious, iterative procedure that provides however reliable estimates. The estimates we present agree, in fact, with the results of the Monte-Carlo simulations performed in a few 1 Hz bands.

In every 1-Hz frequency band, where a CW signal has been found, we decrease the peak amplitude $\mathcal{R}$ by a factor varying from 1 up to 0.1, in steps of 0.1, and select all peaks
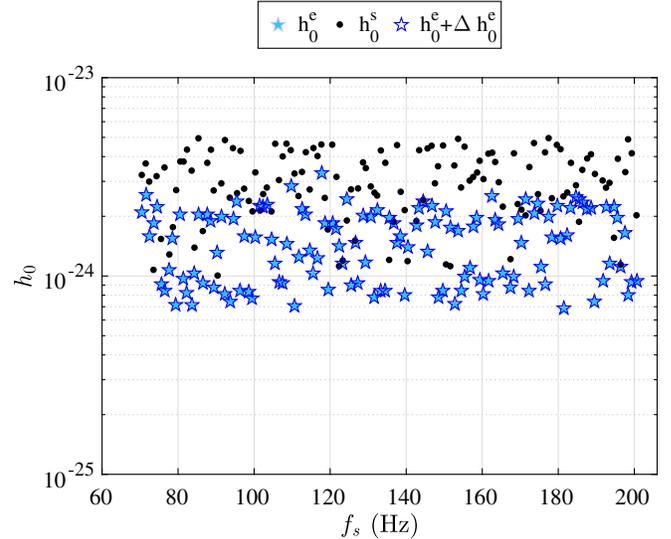


FIG. 9.  Detectable strain amplitudes versus 128 signal frequencies. The dots are the injected strain values, the filled stars are the minimum strain amplitudes that can be detected (over one month of single-detector data), and the (hardly visible) empty stars show the one-sigma errorbars associated to these estimates.

above the initial threshold $\mathcal{R}_{\text{th}} = \sqrt{2.5}$. To these rescaled data we apply, recursively, the cascade of four filters introduced in Sec. IV until a CW signal can no longer be detected. This translates into testing essentially the performance of the Gaussian-like filter, which has the final say-so to consider a signal detected or not detected. Such approach is equivalent to keep fixed all parameters of the sources injected in every 1-Hz band, but the strain amplitude $h_0^s$, which decreases by a factor that can be obtained from Eq. (51). We determine thus $N_p$ and $E_p$, based on the scaled peak amplitude $\mathcal{R}$, and crossing the same threshold (see Appendix D for more details). Hence, we can estimate the minimum strain amplitude $h_0^e$ that can be detected by the filter cascade. The resulting detectable strain amplitude $h_0^e$ is plotted in Fig. 9 against signal frequency for 128 sources, as we excluded the three 1-Hz frequency bands where no signal has been previously detected.

The results shown in Fig. 9 are quite promising, mainly considering they derive from one month of single-detector data, albeit with the caveat of being in Gaussian noise. The broad variability in terms of sensitivity estimation, visible in Fig. 9, is attributed to the performance of the Gaussian-like filter, which needs to be fine-tuned and enhanced. This is part however of a supplementary study, which aims also at carrying investigations over to real interferometer data, with the goal of both testing and strengthen the performance of such a filter, and provide more accurate sensitivity estimations. In the circumstance of real detector data we expect several outliers to compromise the ability of such a filter to identify putative CW signals. We plan however to resort to follow-up studies, and coincidence-based methods, to verify the presence of a given signal [24,59].

## XI. COMPUTING COST MODEL

The computational cost of the machinery here presented is estimated to be around 2.4 CPU hours on a 1.4 GHz Intel Core i5 processor. This estimate is based on timing the different pieces of the analysis (items from 2 to 7 listed in Sec. V), without counting the SFDB production, which is a common part of other CW searches and noise characterization studies, and is discussed in [68].

The total computing cost $C_{tot}$ needed to both claim a detection and estimate source parameters is

$$C_{tot} = C_1 + C_2, \qquad (52)$$

where $C_1$ is a computing cost scaling linearly with the time length (i.e., the observation time) of the data being analyzed, with the number of inspected frequency bands $N_b$, and inversely with the FFT duration. Furthermore, if we would analyze data from different detectors independently, $C_1$ would scale also with the number of interferometers considered ($N_{det}$), i.e.,

$$C_1 = c_0 N_b \mathcal{J} N_{det} \frac{T_{obs}}{T_{FFT}} \mathcal{I}, \qquad (53)$$

where $c_0 \sim 6$ ms is a timed implementation cost per subpeakmap. We note that $T_{obs}/T_{FFT}$ is exactly the number of FFTs $\mathcal{N}$ only for gapless data; $\mathcal{I}$ and $\mathcal{J}$ are factors equal to 2 for interlaced (by half) FFTs and frequency bands, respectively, and 1 otherwise.

In other words, $C_1$ is the time required to perform steps from 2 to 6 listed in Sec. V. The remaining step 7, needed to provide further parameter estimates, is evaluated in a very cheap amount of time expressed by

$$C_2 = c_\theta N_b' \mathcal{J}, \qquad (54)$$

where $c_\theta \sim 2.34$ s is an implementation and hardware-dependent computing cost per band. We note that $N_b' \lesseqqgtr N_b$ is the number of frequency bands where a signal has been identified.

Hence, for $T_{obs} = 1$ month, $T_{FFT} = 512$ s, $N_b = 131$, $\mathcal{J} = 1$, $N_{det} = 1$, $\mathcal{I} = 2$, and $N_b' = 128$, we have $C_1 \sim 2.3$ hours and $C_2 \sim 5$ minutes.

We underline that the search frequency range is relevant to computational performance only because it affects the Doppler shift due to orbital motion. In fact, an increase in frequency would correspond to a reduction in $T_{FFT}$ [see Eq. (B6) for details], and a consequent increase in $C_1$. The method will remain however computationally tractable on a single-processor computer[22]

---

[22]A similar reasoning applies when we reduce the source orbital period and increase the semimajor axis with respect to the choices applied here.

There are other CW pipelines, such as those included in the Scorpius X-1 MDC [36] (in particular the "Radiometer" search, which is roughly comparable to this method in computational efficiency) and [69], to which more detailed comparisons will be made in the future.

## XII. CONCLUSIONS

This paper describes an incoherent and computationally cheap innovative method to search for continuous gravitational waves emitted by pulsars orbiting a companion object. To show the pipeline performance, we analyze one month of simulated gapless Gaussian noise single-detector data to which we added 131 CW signals emitted by pulsars in low- and high-eccentricity binary systems. We used an advanced LIGO-Virgo detector design sensitivity of $S_h = 4 \times 10^{-24}$ Hz$^{-1/2}$[11–13] constant over the search frequency range [70,200] Hz and reported 128 detections, with the weakest—injected and detected—gravitational-wave amplitude of $h_0 \sim 10^{-24}$. We point out that, by using one year of data and three detectors in their advanced configuration, this translates into being able to detect CW signals with strain amplitude as low as $h_0 \sim 3 \times 10^{-25}$. At very small frequencies, this corresponds to reach, and go below, the torque balance limit currently foreseen for Scorpius X-1 [28]. We also stress that, as discussed in Sec. X, depending on the orbital Doppler modulation, our novel algorithm has the ability to detect also signals as low as $h_0 \sim 7 \times 10^{-25}$ for a single-detector one-month data. So, the reachable sensitivity is expected to improve further with one year of three-detector data.

After claiming a detection, we recover the signal parameters, which we infer with decent accuracy. Note that, contrary to the Scorpius X-1 MDC [36] and previous searches [33], in addition to circular orbits, we consider also high-eccentricity orbits, being able to process data on a single-processor computer, and attaining a good sensitivity. A detailed comparison study should be however carried out between the current algorithm and other existing methods, such as [36,69–76], to exactly assess the sensitivity performance and compare the related computational costs. This is, however, beyond the scope of the current paper, and we defer it to future work.

The current search may also be used as a fast quick-look analysis to scan the data and possibly single out significant candidate signals which deserve further investigations. The recovery of source parameters can be honed by assuming the estimates obtained for signal frequency and orbital parameters as preliminary. Hence, one can envisage using those estimates to approximately demodulate the data from the binary orbital motion. This will serve to create longer FFTs, enhancing, thus, the search sensitivity. Rather than considering a fixed FFT duration for the entire search frequency range (as done here), we note that the sensitivity can be further increased by creating FFTs of different length in time, based on the highest frequency value of the

subinterval into which the whole search frequency range has been previously split. This will allow a higher sensitivity at low frequencies. At this point, the procedure outlined here can be reiterated to get more accurate signal parameter estimates. Alternatively, or in addition, deep follow-up studies with a longer coherence time for the weakest detections can be applied to increase the SNR of a putative signal, and obtain even more precise parameter estimates (e.g., [24,59]). This entails the generalization of the resampling technique, currently used for CW signals from isolated pulsars [24], to the case of binary systems, and this is part of a separate ongoing study.

Although the current work is invaluable to both gauge and validate the performance of the various steps of the presented algorithm, accurate testing is needed (and currently underway) to enhance the Gaussian-like filter rendering and efficiency. In fact, the success of such a filter varies depending both on the particular source orbital parameters, which govern the shape of the modulated signal pattern, and on the signal strength.

Looking farther forward, we expect to generalize the method presented here to search for a broader class of signals, considering spindown parameters and accretion-induced spin-wandering effects for the simulated sources, which will be added to real data collected from multiple detectors. We expect this novel methodology to be so robust not to be subject to frequency variabilities, and so the characteristics of spin wandering are expected to have no impact on the analysis. The complication to deal with real detector data will be faced applying *ad hoc* noise reduction approaches that aim to identify, and possibly knock out, non-Gaussian artifacts, as well as coincidence-based approaches [24,77].

Furthermore, due to the tremendously cheap 2.4 CPU hours taken to analyze a fixed sky location, 131 1-Hz frequency bands, and one month of single-detector data, the current procedure can be applied in all-sky schemes in an equally successful and inexpensive way. This can be in fact achieved analyzing in parallel as many sky positions as permitted by the available computational power.

In addition, we plan to improve the current pipeline in order to provide estimates of source orientation and polarization parameters (i.e., $\cos \iota$, $\psi$, and $\phi_0$), and also to get further enhancements in strain sensitivity, which will allow us to detect signal amplitudes at, or below, the torque balance limit currently foreseen for Scorpius X-1-like sources.

The present method—and its future improvements—will be applied to analyze new data collected by the ever-sensitive Advanced LIGO and Virgo detectors (which are second-generation detectors), as well as by future third-generation detectors. This is an outstanding challenge, which makes us more optimistic about being able to make direct detections of CW signals. Such detections will provide new insights into the internal structure, formation history and population statistics of neutron stars. In case of no detection, we plan however to set more astrophysically constraining upper limits on the gravitational-wave signal strength.

## APPENDIX A: MAXIMAL DOPPLER SHIFT DUE TO ORBITAL MOTION

As shown in [28], we estimate the maximal Doppler shift the intrinsic signal frequency of a binary CW signal can undergo due to orbital motion. From the approximated phase model of Eq. (13), we see that the instantaneous Doppler shift is

$$
\left| \frac{\dot{\phi}(t)}{2\pi f} - 1 \right| = \left| \frac{\dot{R}(t)}{c} \right|
$$
$$
= a_p \Omega \left| \frac{\sqrt{1 - e^2} \cos E \cos \omega - \sin E \sin \omega}{1 - e \cos E} \right|
$$
$$
\leq a_p \Omega \frac{|\sin E \sin \omega| + |\cos E \cos \omega|}{|1 - e \cos E|}, \quad \text{(A1)}
$$

where we used the fact that $|a + b| \leq |a| + |b|$ and $\sqrt{1 - e^2} \leq 1$. In addition, we observe that

$$
|\cos E \cos \omega| + |\sin E \sin \omega|
$$
$$
= \max\{|\cos(E + \omega)|, |\cos(E - \omega)|\} \leq 1, \quad \text{(A2)}
$$

and $|1 - e \cos E| \geq 1 - e$ to obtain

$$
\left| \frac{\dot{\phi}(t)}{2\pi f} - 1 \right| = \left| \frac{\dot{R}(t)}{c} \right| \leq \Delta M, \quad \text{(A3)}
$$

with

$$
\Delta M = \frac{a_p \Omega}{1 - e} \quad \text{(A4)}
$$

being the maximal Doppler modulation due to orbital motion.

## APPENDIX B: MAXIMAL FFT DURATION

The maximal length of the FFT is limited by the linear-phase approximation of Eq. (12). In order to improve sensitivity, we want to choose the longest possible FFT duration $T_{\text{FFT}}$ with an acceptable error in the linear-phase approximation. In order to estimate the maximal value of

this phase-error ($|\Delta\phi|$) over a single FFT, we follow [28] and estimate it as

$$|\Delta\phi| = \left| \frac{1}{2}\ddot{\phi}(t)\left(\frac{T_{\text{FFT}}}{2}\right)^2 \right|, \qquad \text{(B1)}$$

but taking the second time derivative of the more precise phase of Eq. (13) rather than that of Eq. (21), as instead done in [28]. Hence, we have

$$\ddot{\phi}(t) = \frac{2\pi f a_p \Omega^2}{(1 - e\cos E)^3}[(\cos E - e)\sin\omega$$
$$+ \cos\omega\sin E\sqrt{1 - e^2}], \qquad \text{(B2)}$$

and replacing Eq. (B2) in Eq. (B1) we obtain the following upper limit on the error in the linear-phase approximation:

$$|\Delta\phi| \le \frac{\pi}{4}T_{\text{FFT}}^2 f a_p \Omega^2 \frac{1+e}{(1-e)^3}, \qquad \text{(B3)}$$

where we used the fact that $|a + b| \le |a| + |b|$, $\sqrt{1 - e^2} \le 1$, and

$$|\cos E\sin\omega| + |\cos\omega\sin E|$$
$$= \max\{|\sin(E+\omega)|, |\sin(E-\omega)|\} \le 1, \qquad \text{(B4)}$$

which brings us to

$$|(\cos E - e)\sin\omega + \cos\omega\sin E\sqrt{1 - e^2}| \le 1 + e. \qquad \text{(B5)}$$

Hence, the FFT length must be

$$T_{\text{FFT}} \sim \frac{2}{\Omega}(1 - e)\sqrt{\frac{1-e}{1+e}\frac{\Delta\phi}{\pi f a_p}}. \qquad \text{(B6)}$$

## APPENDIX C: CONSTRAINING THE FFT DURATION

The time baseline of an FFT is typically determined by requiring the signal power to be concentrated in less than a frequency bin. In Appendix B, we have nailed down the longest possible FFT duration based on the phase error of the linear-phase approximation over an FFT. The estimate provided in Eq. (B6) is, however, very conservative. Indeed, assuming a phase error $\Delta\phi = \pi/4$, and the most unfavorable boundaries of the investigated parameter space, i.e., $P = 10\,$h, $e = 0.9$, $a_p = 3\,$s, $f = 200\,$Hz, we obtain $T_{\text{FFT}} \sim 129\,$s. The FFT durations we used to process the set of data containing 131 sources, simulated with parameters within the ranges provided in Eq. (46), are instead 512 s, which are based on an approximation that

allows us to improve the sensitivity by a factor of $\sqrt{512/129} \sim 2$.

The reasoning behind the usage of FFTs longer than those obtained with Eq. (B6) consists of considering, rather than the whole orbit swept by a source, only a large fraction of it, i.e., at most the 80%. This is somewhat reasonable as, over one month, we will observe from a minimum of 15 orbits to a maximum of 72 orbits (considering $10\,$h $\le P \le 2\,$d). We explain better how we achieved such a consideration. We performed simulations for 1000 sources, with random parameters drawn uniformly from the ranges of Eq. (46), and for each source we compute the rate at which the source changes its velocity by taking the second time derivative of Eq. (6) (i.e., the acceleration divided by the light speed). We then choose, for every source, the 80th percentile of $|\ddot{R}(t)|/c$, neglecting thus the 20% highest values in terms of $|\ddot{R}(t)|/c$, which correspond to parts of the orbit where there is the largest source velocity variation, and then the largest orbital Doppler modulation.

Figure 10 shows the distribution of the 80th percentiles of $|\ddot{R}(t)|/c$ for 1000 simulated sources. We see that the maximum value of the 80th percentiles of $|\ddot{R}(t)|/c$ is $\mathcal{P} \sim 8 \times 10^{-8}\,$s$^{-1}$, which is of course the 100th percentile of the distribution in Fig. 10. We should use such value to reliably estimate the maximum FFT duration as a function of the search frequency:

$$T_{\text{FFT}}^{\text{max}} = \frac{1}{\sqrt{\mathcal{P}f}}. \qquad \text{(C1)}$$

If we consider instead the 85th percentile (marked by the arrow in Fig. 10,) i.e., $\mathcal{P}_{0.85} \sim 2 \times 10^{-8}\,$s$^{-1}$, we can further
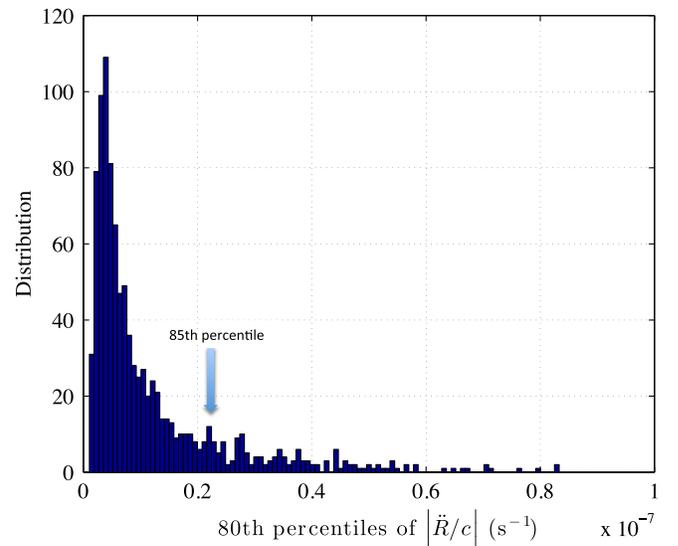


FIG. 10. Distribution of the 80th percentiles of $|\ddot{R}(t)|/c$ for 1000 sources. The arrow denotes the value of $|\ddot{R}(t)|/c$ corresponding to the 85th percentile, i.e., $|\ddot{R}_0.85|/c \sim 2 \times 10^{-8}\,$s$^{-1}$.
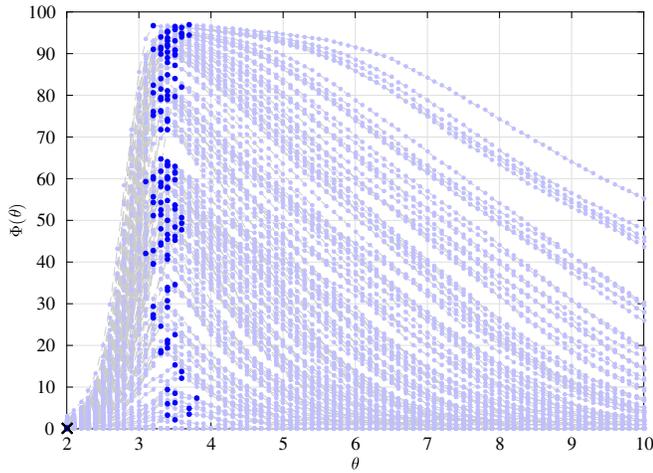
FIG. 11.   Observable $\Phi$ as a function of possible thresholds $\theta$ for 131 analyzed frequency bands. The dark-blue dots correspond to the maximum values of $\Phi(\theta)$, while the black cross at $\theta = 2$ corresponds to frequency bands where no signals could be identified, i.e.: [141,142] Hz, [166,167] Hz, and [190,191] Hz (see details in Sec. VIII).

lengthen the FFT duration by a factor of 2 at the expense of further reducing the fraction of the orbit for 15% of the sources, considering for them even smaller velocity variations.

Hence, replacing $\mathcal{P}$ with $\mathcal{P}_{0.85}$ in Eq. (C1), we obtain $T_{\mathrm{FFT}}^{\max} = 512$ s, which we use to produce FFTs for the whole search frequency range $f \in [70, 200]$ Hz. Such a choice is actually based on the (worst) highest search frequency, i.e., 200 Hz, as for 70 Hz an FFT of 845 s would be more appropriate. As this translates into a reduction in sensitivity (and then in parameter estimate accuracy), one can envisage to produce SFDB of different duration, depending on the highest frequency value of the subinterval into which the entire search frequency range has been previously split.

## APPENDIX D: SETTING A STRINGENT THRESHOLD FOR PEAK SELECTION

The threshold $\theta_{\mathrm{thr}}$, which we choose in order to select the most significant peaks, affects the parameter estimate abilities and search sensitivity. The criterion we use for the choice of $\theta_{\mathrm{thr}}$ is the maximization of an informative observable $\Phi(\theta)$, which combines the number of peaks solely due to signal, $N_s(\theta)$, and the number of peaks due to pure noise, $N_n(\theta)$, as a function of a varying threshold $\theta$, and is empirically found to be

$$\Phi(\theta) = \frac{N_s(\theta)}{\sqrt{N_s(\theta) + N_n(\theta)^2}}. \tag{D1}$$

We identify frequency subbands, containing the CW signature, by using the procedure described in Sec. IV. We then select peaks with an amplitude $\mathcal{R}$ larger than $\theta$, ranging from 2 to 10 in steps of 0.1; we note that the minimum value of $\theta$ is chosen to be slightly larger than the threshold set for initial peak selection $\mathcal{R}_{\mathrm{th}} = \sqrt{2.5} \sim 1.6$.

Then, we can estimate the number of peaks due to signal and noise in that band, able to surpass a given value of $\theta$, i.e., $N_{\mathrm{sn}}(\theta)$ [blue dots in Fig. 2(a)], and the number of peaks due to pure noise in the same band, but assuming no signals were present, and surpassing the same $\theta$, i.e., $N_n(\theta)$ [red dots in Fig. 2(a)]. Hence, we compute the number of peaks solely due to signal in a certain frequency band as $N_s(\theta) = N_{\mathrm{sn}}(\theta) - N_n(\theta)$. The observable $\Phi$, as a function of $\theta$, is shown in Fig. 11.

As the threshold in a given frequency subband, we choose the optimal value, which is the one maximizing $\Phi(\theta)$. As shown in Fig. 11, this varies from a minimum of 3.1 to a maximum of 3.8, with an average of $\theta_{\mathrm{thr}} = 3.5$ for the bulk of frequency bands. In every subband, where a presumed CW signal has been identified, we select all



FIG. 12.   $\mathcal{G}$ function of Eq. (25) versus $n_t$ for $z = 2$, $u = 2$, $q = 1.7$, $g = 10$ and three $\sigma$ values [panel (a)], and $\sigma = 6$ with different values of $z$, $u$, $q$, $g$, shown in the legend [panel (b)].

peaks above this slightly varying threshold, and average the related frequencies corresponding to a same FFT midtime, obtaining at most $\mathcal{N}$ $(t_c, \bar{f}_p)$ pairs.

A systematic study of how varies $\theta_{\text{thr}}$ with varying $T_{\text{obs}}$ and $T_{\text{FFT}}$ lies outside the scope of this paper as a value of $\theta_{\text{thr}} = 3.5$ can be broadly adopted for further analyses, even on real detector data. Such a value is indeed consistent with what can be obtained by a visual inspection of the analyzed frequency bands, and is however at the edge of what would be indistinguishable from Gaussian noise [as can be appreciated in Fig. 2(a)].

## APPENDIX E: GAUSSIAN-LIKE FILTER DETAILS

The $\mathcal{G}$ function of Eq. (25) is plotted in Fig. 12(a) for different $\sigma$ (i.e., "standard deviation") values and $z = 2$, $u = 2$, $q = 1.7$, $g = 10$, while in Fig. 12(b) for different values of $z$, $u$, $q$, $g$ and $\sigma = 6$.

Simulation-based studies bring us to choose a $\mathcal{G}$ function with $z = 2$, $u = 2$, $q = 1.7$, $g = 10$, and $\sigma = 6$ (red curve in Fig. 12).

We stress that the criterion given in Eq. (23) is not sufficient to prevent the selection of instrumental artifacts, but it is used to quickly sift the analyzed frequency bands, and identify the disturbed ones, which will be inspected with more scrutiny. Furthermore, in the present work we have simulated a single detector data set adding fake signals into Gaussian noise data, but when applying the current method to real case, data taken from multiple detectors will be available, allowing us to more reliably exclude prominent disturbances by using a coincidence analysis. The application of more aggressive noise identification and artifact mitigation techniques is also envisaged to be beneficial, and is currently under investigation.

The false alarm probability for the identification of a signal (i.e., a false alarm) in frequency bands where no signal is present is computed generating $6 \times 10^4$ pure Gaussian noise realizations for which we verify if the condition expressed by Eq. (23) is satisfied. Since this is not fulfilled, we can place an upper limit on the false alarm probability, which results being smaller than $1.6 \times 10^{-5}$. We note that such a value is very conservative as it does not take into account that a false alarm, if found, will undergo further checks via the Gaussian-like filter before being claimed a detection.

## APPENDIX F: EXAMPLE OF PERIODOGRAM APPEARANCE

Figure 13 shows the periodograms evaluated for $(t_c, \bar{f}_p)$ pairs selected from randomly generated pure Gaussian noise [panel (a)], and from data where two CW signals have been added to artificial gapless Gaussian noise, with orbital eccentricity $e = 0$ and $e \sim 0.87$ for the panels (b)
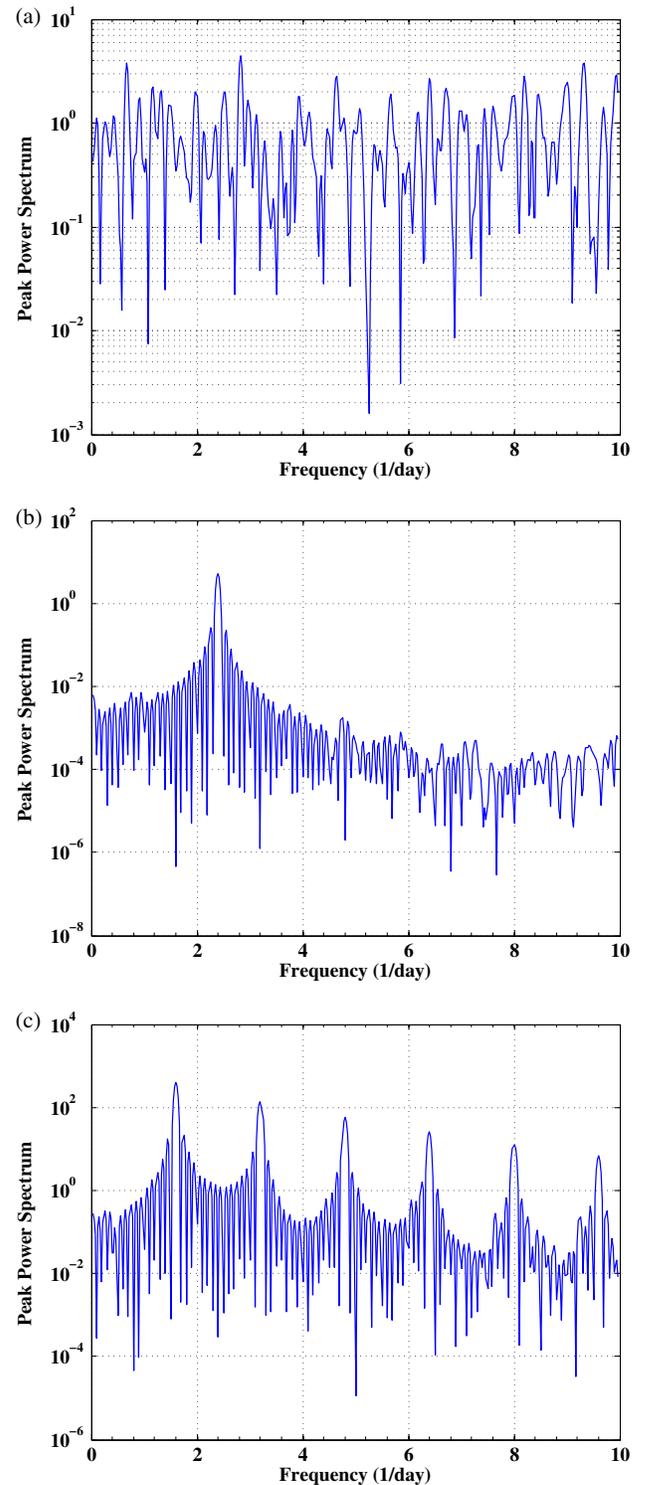


FIG. 13.   Semilogarithmic plots of the periodogram as a function of the frequency for: normally distributed random numbers (a); set of synthetic data where two CW signals have been added at 155.5 Hz (b), and 1001.5 Hz (c), with orbital eccentricities $e = 0$ (b), and $e \sim 0.87$ (c), and strain amplitude $h_0 \sim 10^{-21}$ (b), and $h_0 = 4 \times 10^{-20}$ (c), respectively. The power spectral density used for these tests is $S_h = 4 \times 10^{-24}$ Hz$^{-1/2}$. The $x$ axis shows, for all panels, 400 frequencies $\nu_j$ at which the periodogram of Eq. (27) has been evaluated for $T_{\text{obs}} = 10$ days and $r_\nu = 4$.

and (c), respectively. The artificially generated data sets span a period of $T_{\mathrm{obs}} = 10$ days, and the conservative FFT duration used is $T_{\mathrm{FFT}} = 128$ s.

A strong component at $\nu \sim 2.4$ d$^{-1}$ is clearly observed in the periodogram of Fig. 13(b), indicating the presence of a signal with orbital period $P = \nu^{-1} \sim 10$ h. The signal frequency is $\sim155.5$ Hz. In Fig. 13(c) we can appreciate the harmonics due to the orbital eccentricity, which is $e \sim 0.87$. The first harmonic at $\nu_1 \sim 1.6$ d$^{-1}$ is the fundamental, and the other harmonics are its multiples. All the harmonics are separated in frequency by $1/P$. The recovered orbital period is $P = \nu_1^{-1} \sim 15$ h. This simulated signal has a frequency of 1001.5 Hz.

We emphasize that both the power spectrum estimate, given by Eq. (27), and the evaluation of the Lomb-Scargle periodogram bring us to the same results.

## APPENDIX G: HARMONIC CONTENT OF THE ORBITAL DOPPLER MODULATION

We estimate the harmonic content of the orbital Doppler modulation $\dot{R}(t)$ (i.e., the amplitudes of the fundamental and its multiples) in order to understand what are the harmonics that contribute to estimate the orbital eccentricity.

The time derivative of the Rømer delay in Eq. (6) is given by

$$\frac{\dot{R}(t)}{c} = a_p \dot{E}(t) \times [-\sin\omega \sin E(t) + \cos\omega \cos E(t) \sqrt{1-e^2}],$$

$$(G1)$$

where $\dot{E}(t) = \Omega/[1 - e\cos E(t)]$ is obtained deriving Eq. (7) with respect to time.

In order to perform a spectral analysis of $\dot{R}(t)/c$, we first find the eccentric anomaly $E$ by numerically solving
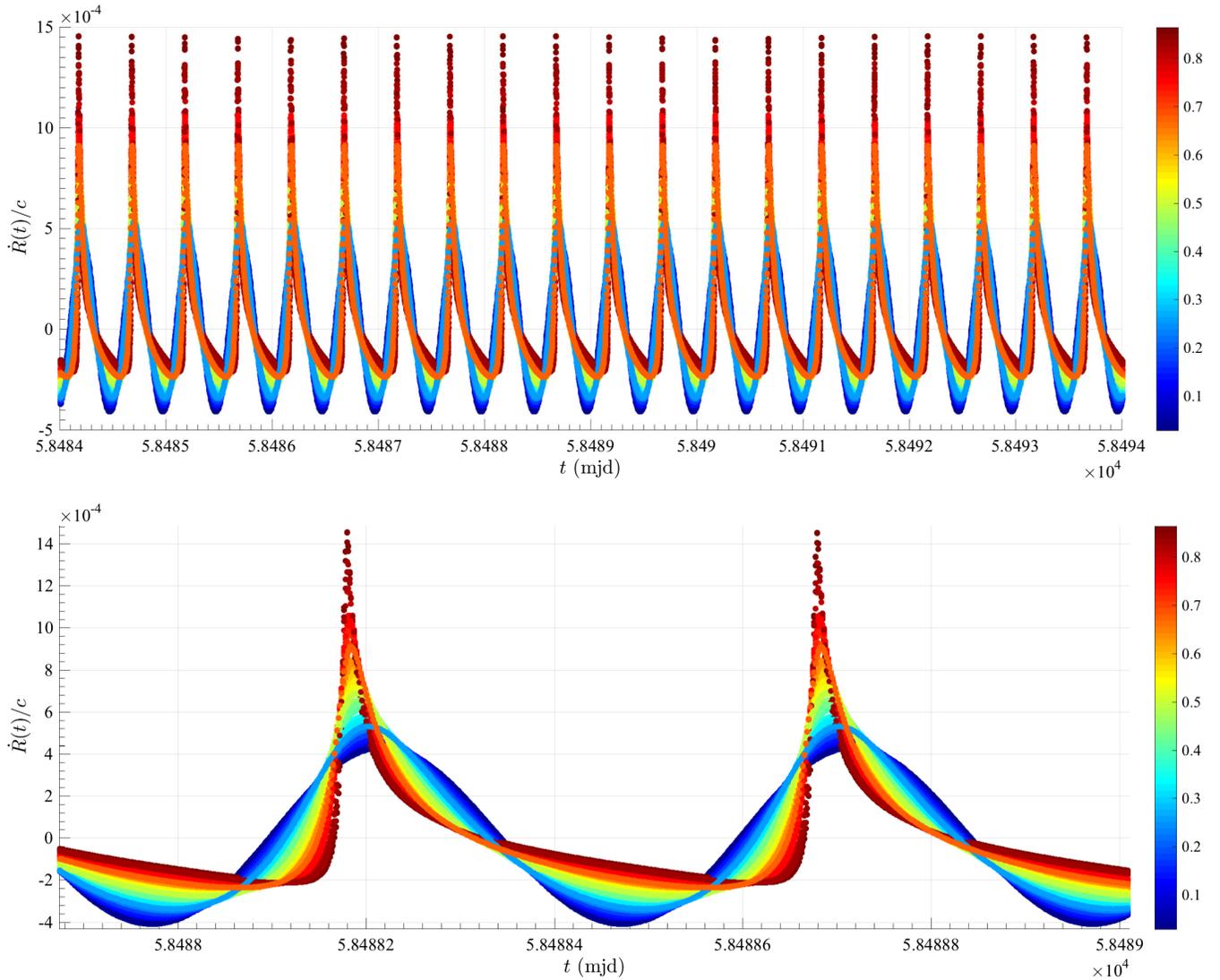


FIG. 14.   Top panel: Time derivative of the Rømer delay, $\dot{R}(t)/c$ [given by Eq. (G1)], versus time for 60 simulated sources with fixed values of $a_p \sim 3$ s, $P \sim 0.5$ d, $\omega \sim 6$ rad, $t_p \sim 54131$ mjd, but $e$ randomly drawn from a uniform distribution in the range $e \in [10^{-5}, 0.9]$, indicated by the values on the colour bar. Bottom panel: The same as top panel, but zooming into roughly a single period.

Eq. (7) and inserting it in Eq. (G1). We consider 60 fixed values of $a_p \sim 3$ s, $P \sim 0.5$ d, $\omega \sim 6$ rad, $t_p \sim 54131$ mjd, and we draw $e$ from a uniform distribution in the range $e \in [10^{-5}, 0.9]$. Equation (7) is solved by employing standard iterative methods, such as the Newton's method and, in case of failure, the bisection method. In Fig. 14 we plot the resulting $\dot{R}(t)/c$ as a function of time. From the top panel we observe, as expected, that each peak is separated from the next by the chosen orbital period $P \sim 0.5$ d, and that sharp-edge peaks correspond to high eccentricity values (as shown on the vertical colour bar).

Given the periodic nature of $\dot{R}(t)/c$ (see Fig. 14), we perform a Fourier decomposition of $\dot{R}(t)/c$ (by computing a discrete Fourier transform), and we plot in Fig. 15 the normalized harmonic amplitude for the first 10 harmonics of $\dot{R}(t)/c$ as a function of the eccentricity, and for 60 values of $\{a_p, t_p, P, e, \omega\}$. We stress that considering more points would only thicken the curves shown in Fig. 15, without adding further information. Independently of the chosen values for $a_p, P, \omega, t_p$, we note that, in order to estimate the orbital eccentricity $e$, only the first two harmonics of $\dot{R}(t)/c$ are necessary, whose contribution is the most dominant.

We further note that for very low-eccentricity orbits, only the first harmonic exists, which is not enough to precisely recover eccentricity values, which cannot then be distinguishable from zero, as stated in Sec. VIII. This is confirmed by a separate investigation of the very low range $e \in [10^{-8}, 10^{-3}]$.

From Fig. 15, we observe that the higher the eccentricity, the greater the number of contributing harmonics.

Furthermore, the amplitude of higher ($\geq 10$) order harmonics is low enough that they can be entirely ignored.

## APPENDIX H: ESTABLISHING THE EXISTENCE OF THE PEAK POWER SPECTRUM HARMONICS

The peak power spectrum described in Sec. VI A 1 is used to derive the source orbital period, which is given by the reciprocal of the fundamental frequency $\nu^m$, i.e., the first
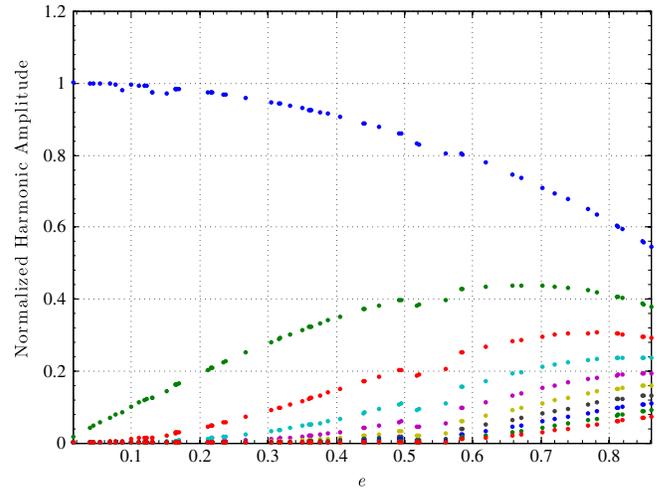


FIG. 15. Normalized harmonic amplitude for the first 10 harmonics of $\dot{R}(t)/c$, and for 60 artificial sources with eccentricity $e \in [10^{-5}, 0.9]$, and fixed $a_p, P, \omega, t_p$ (see main text for details).

harmonic [see Eq. (29)]. The second harmonic (i.e., $2\nu^m$), when exists, is instead used to determine the source orbital eccentricity and argument of periapse [see Eqs. (42) and (43), respectively]. In order to check if these harmonics exist in the peak power spectrum, we use a threshold based on a robust estimator, such as the median. After identifying $\nu^m$ and $2\nu^m$ in the peak power spectrum, we compute the median and dispersion parameter for roughly 80 samples around those two frequencies, and verify if the amplitude of each harmonic satisfy the following condition,

$$\mathcal{H}_h > \mathcal{M}(1) + \mathcal{M}(2), \tag{H1}$$

with $h = 1, 2$, $\mathcal{M}(1) = \text{median}(\mathcal{S}(h\nu^m - 40d\nu : h\nu^m + 40d\nu))$, and $\mathcal{M}(2) = \text{median}(|\mathcal{S}(h\nu^m - 40d\nu : h\nu^m + 40d\nu) - \mathcal{M}(1)|)$.

[1] J. Weber, Phys. Rev. **117**, 306 (1960).

[2] P. Astone *et al.* (International Gravitational Event Collaboration), Phys. Rev. D **68**, 022001 (2003).

[3] P. Leaci and A. Ortolan, Phys. Rev. A **76**, 062101 (2007).

[4] P. Leaci, A. Vinante, M. Bonaldi, P. Falferi, A. Pontin, G. A. Prodi, and J. P. Zendri, Phys. Rev. D **77**, 062001 (2008).

[5] P. Leaci, T. D. Canton, A. Ortolan, and G. A Prodi, Classical Quantum Gravity **25**, 195018 (2008).

[6] P. Astone et al., Phys. Rev. D **82**, 022003 (2010).

[7] M. Ando *et al.* (TAMA Collaboration), Phys. Rev. Lett. **86**, 3950 (2001).

[8] H. Grote and (The LIGO Scientific Collaboration), Classical Quantum Gravity **27**, 084003 (2010).

[9] T. Accadia *et al.*, J. Inst. **7**, P03012 (2012).

[10] B. P. Abbott *et al.* (The LIGO Scientific Collaboration), Rep. Prog. Phys. **72**, 076901 (2009).

[11] J. Aasi *et al.* (The LIGO Scientific Collaboration), Classical Quantum Gravity **32**, 074001 (2015).

[12] F. Acernese *et al.* (The VIRGO Collaboration), Classical Quantum Gravity **32**, 024001 (2015).

[13] B. P. Abbott *et al.* (The VIRGO Collaboration and The LIGO Scientific Collaboration), Living Rev. Relativ. **19**, 1 (2016).

[14] B. P. Abbott *et al.* (The VIRGO Collaboration and The LIGO Scientific Collaboration), Phys. Rev. Lett. **116**, 131103 (2016).
[15] B. P. Abbott *et al.* (The LIGO Scientific Collaboration and The VIRGO Collaboration), Phys. Rev. Lett **116**, 061102 (2016).
[16] B. J. Owen, Phys. Rev. Lett. **95**, 211101 (2005).
[17] M. Camenzind, *Compact Objects in Astrophysics: White Dwarfs, Neutron Stars and Black Holes* (Springer, New York, 2007).
[18] ATNF catalogue, http://www.atnf.csiro. au/people/pulsar/psrcat/.
[19] P. D. Lasky, Pub. Astron. Soc. Aust. **32**, e034 (2015).
[20] S. V. Dhurandhar and A. Vecchio, Phys. Rev. D **63**, 122001 (2001).
[21] P. Jaranowski, A. Krolak, and B. F. Schutz, Phys. Rev. D **58**, 063001 (1998).
[22] P. Astone *et al.*, Phys. Rev. D **65**, 022001 (2001).
[23] J. Abadie *et al.* (The VIRGO Collaboration and The LIGO Scientific Collaboration), Astrophys. J. **737**, 93 (2011).
[24] J. Aasi *et al.* (The VIRGO Collaboration and The LIGO Scientific Collaboration), Phys. Rev. D **91**, 022004 (2015).
[25] P. R. Brady and T. Creighton, Phys. Rev. D **61**, 082001 (2000).
[26] B. Krishnan, A. M. Sintes, M. A. Papa, B. F. Schutz, S. Frasca, and C. Palomba, Phys. Rev. D **70**, 082001 (2004).
[27] C. Cutler, I. Gholami, and B. Krishnan, Phys. Rev. D **72**, 042004 (2005).
[28] P. Leaci and R. Prix, Phys. Rev. D **91**, 102003 (2015).
[29] J. Aasi *et al.* (The VIRGO Collaboration and The LIGO Scientific Collaboration), Phys. Rev. D **87**, 042001 (2013).
[30] J. Aasi *et al.* (The VIRGO Collaboration and The LIGO Scientific Collaboration), Phys. Rev. D **93**, 042007 (2016).
[31] J. Aasi *et al.* (The VIRGO Collaboration and The LIGO Scientific Collaboration), Phys. Rev. D **88**, 102002 (2013).
[32] J. Aasi *et al.* (The VIRGO Collaboration and The LIGO Scientific Collaboration), Phys. Rev. D **90**, 062010 (2014).
[33] J. Aasi *et al.* (The VIRGO Collaboration and The LIGO Scientific Collaboration), Phys. Rev. D **91**, 062008 (2015).
[34] A. Watts, B. Krishnan, L. Bildsten, and B. F. Schutz, Mon. Not. R. Astron. Soc. **389**, 839 (2008).
[35] B. Abbott *et al.* (The LIGO Scientific Collaboration), Phys. Rev. D **76**, 082001 (2007).
[36] C. Messenger *et al.*, Phys. Rev. D **92**, 023006 (2015).
[37] R. A. Hulse and J. H. Taylor, Astrophys. J. **195**, L51 (1975).
[38] A. L. Watts *et al.*, Rev. Mod. Phys. **88**, 021001 (2016).
[39] R. V. Wagoner, Astrophys. J. **278**, 345 (1984).
[40] L. Bildsten, Astrophys. J. **501**, L89 (1998).
[41] D. Chakrabarty, E. H. Morgan, M. P. Muno, D. K. Galloway, R. Wijnands, M. van der Klis, and C. B. Markwardt, Nature (London) **424**, 42 (2003).
[42] G. Cook, S. Shapiro, and S. Teukolsky, Astrophys. J. **424**, 823 (1994).
[43] B. Haskell *et al.*, Proceedings, 3rd Session of the Sant Cugat Forum on Astrophysics: Gravitational Waves Astrophysics: Sant Cugat, Catalonia, Spain, April 22-25, 2014, *Astrophys. Space Sci. Proc.* 40, 85 (2015).
[44] S. Bonazzola and E. Gourgoulhon, Astron. Astrophys. **312**, 675 (1996).
[45] P. Leaci, J. Phys. Conf. Ser. **354**, 012010 (2012).
[46] G. Ushomirsky, C. Cutler, and L. Bildsten, Mon. Not. R. Astron. Soc. **319**, 902 (2000).
[47] R. Smoluchowski, Phys. Rev. Lett. **24**, 923 (1970).
[48] M. Ruderman, Astrophys. J. **382**, 587 (1991).
[49] C. Horowitz and K. Kadau, Phys. Rev. Lett. **102**, 191102 (2009).
[50] J. Papaloizou and J. E. Pringle, Mon. Not. R. Astron. Soc. **184**, 501 (1978).
[51] G. B. Hobbs, R. T. Edwards, and R. N. Manchester, Mon. Not. R. Astron. Soc. **369**, 655 (2006).
[52] R. T. Edwards, G. B. Hobbs, and R. N. Manchester, Mon. Not. Roy. Astron. Soc. **372**, 1549 (2006).
[53] P. Astone, S. Frasca, and C. Palomba, Gravitational wave data analysis. Proceedings, 9th Workshop, GWDAW 2004, Annecy, France, December 15-18, 2004, *Classical Quantum Gravity* 22, S1197 (2005).
[54] P. Astone, A. Colla, S. D'Antonio, S. Frasca, and C. Palomba, Phys. Rev. D **90**, 042002 (2014).
[55] F. Acernese *et al.*, Proceedings, 13th Gravitational Wave Data Analysis Workshop on Bridging gravitational wave astronomy and observational astrophysics (GWDAW13), *Classical Quantum Gravity* 26, 204002 (2009).
[56] P. Leaci, P. Astone, M. A. Papa, and S. Frasca, J. Phys. Conf. Ser. **228**, 012006 (2010).
[57] B. Abbott *et al.* (The LIGO Scientific Collaboration), Phys. Rev. D **77**, 022001 (2008); **80**, 129904(E) (2009).
[58] SNAG v2, http://grwavsf.roma1.infn.it/snag/.
[59] M. Shaltev, P. Leaci, M. A. Papa, and R. Prix, Phys. Rev. D **89**, 124030 (2014).
[60] E. M. Standish, Astron. Astrophys. **114**, 297 (1982).
[61] E. M. Standish, Astron. Astrophys. **233**, 252 (1990).
[62] F. Eng, F. Gunnarsson, and F. Gustafsson, IEEE Trans. Signal Processing **56**, 3089 (2008).
[63] J. D. Scargle, Astrophys. J. **263**, 835 (1982).
[64] N. R. Lomb, Astrophys. Space Sci. **39**, 447 (1976).
[65] A. G. Barnett and A. J. Dobson, *Analysing Seasonal Health Data* (Springer, Berlin, 2010).
[66] LALSUITE, https://www.lsc-group.phys.uwm.edu/daswg/projects/lalsuite.html, (Version 9fc768e5c4e5439-b5a8e3c5f68805539a23e5d05).
[67] R. N. Manchester, G. B. Hobbs, A. Teoh, and M. Hobbs, Astron. J. **129**, 1993 (2005).
[68] The AdV Computing Model-VIR-0129H-13 (The VIRGO Collaboration), https://tds.virgo-gw.eu/?content=3&r=10776.
[69] S. Suvorova, L. Sun, A. Melatos, W. Moran, and R. J. Evans, Phys. Rev. D **93**, 123009 (2016).
[70] S. W. Ballmer, Gravitational waves. Proceedings, 6th Edoardo Amaldi Conference, Amaldi6, Bankoku Shinryou-kan, June 20-24, 2005, *Classical Quantum Gravity* 23, S179 (2006).
[71] C. Messenger and G. Woan, Classical Quantum Gravity **24**, S469 (2007).
[72] L. Sammut, C. Messenger, A. Melatos, and B. J. Owen, Phys. Rev. D **89**, 043001 (2014).
[73] E. Goetz and K. Riles, Classical Quantum Gravity **28**, 215006 (2011).
[74] G. D. Meadors, E. Goetz, and K. Riles, Classical Quantum Gravity **33**, 105017 (2016).
[75] S. Dhurandhar, B. Krishnan, H. Mukhopadhyay, and J. T. Whelan, Phys. Rev. D **77**, 082001 (2008).
[76] J. T. Whelan, S. Sundaresan, Y. Zhang, and P. Peiris, Phys. Rev. D **91**, 102005 (2015).
[77] P. Leaci, Phys. Scr. **90**, 125001 (2015).