Optimally setting up directed searches for continuous gravitational waves in Advanced LIGO O1 data

Jing Ming, Maria Alessandra Papa, Badri Krishnan, Reinhard Prix, Christian Beer, Sylvia J. Zhu, Heinz-Bernd Eggenstein, Oliver Bock, and Bernd Machenschalk

Phys. Rev. D **97**, 024051 — Published 31 January 2018

DOI: 10.1103/PhysRevD.97.024051

Optimally setting up directed searches for continuous gravitational waves in Advanced LIGO O1 data

Jing Ming[1,2,3], Maria Alessandra Papa[1,2,4], Badri Krishnan[2,3], Reinhard Prix[2,3], Christian Beer[2,3], Sylvia J. Zhu[1,2], Heinz-Bernd Eggenstein[1,2], Oliver Bock[2,3], Bernd Machenschalk[2,3]
[1] *Max-Planck-Institut für Gravitationsphysik, Albert Einstein Institute, am Mühlenberg 1, 14476, Potsdam-Golm, Germany*
[2] *Max-Planck-Institut für Gravitationsphysik, Albert Einstein Institute, Callinstraβe 38, 30167, Hannover, Germany*
[3] *Leibniz Universität Hannover, Welfengarten 1, 30167, Hannover, Germany*
[4] *University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201, USA*

In this paper we design a search for continuous gravitational waves from three supernova remnants: Vela Jr., Cassiopeia A (Cas A) and G347.3. These systems might harbor rapidly rotating neutron stars emitting quasi-periodic gravitational radiation detectable by the advanced LIGO detectors. Our search is designed to use the volunteer computing project Einstein@Home for a few months and assumes the sensitivity and duty cycles of the advanced LIGO detectors during their first science run. For all three supernova remnants, the sky-positions of their central compact objects are well known but the frequency and spin-down rates of the neutron stars are unknown which makes the searches computationally limited. In a previous paper we have proposed a general framework for deciding on what target we should spend computational resources and in what proportion, what frequency and spin-down ranges we should search for every target, and with what search set-up. Here we further expand this framework and apply it to design a search directed at detecting continuous gravitational wave signals from the most promising three supernova remnants identified as such in the previous work. Our optimization procedure yields broad frequency and spin-down searches for all three objects, at an unprecedented level of sensitivity: The smallest detectable gravitational wave strain h_0 for Cas A is expected to be 2 times smaller than the most sensitive upper-limits published to date, and our proposed search, which was set-up and ran on the volunteer computing project Einstein@Home, covers a much larger frequency range.

I. INTRODUCTION

Continuous gravitational wave (CW) signals at frequencies between $\approx 10 - 1000$ Hz are expected to be emitted by rapidly rotating compact objects with shapes that are not perfectly axially symmetric. Because these signals are extremely weak [1–3], one needs to combine the data over very long periods of time (months) in order to raise the signal significantly above the average noise level. On the other hand the ability to resolve different waveforms increases very quickly with the duration of the observation time, and hence CW searches over broad ranges of different waveforms are computationally very expensive. We can broadly characterize CW searches as targeted, all-sky and directed according to the computational cost requirements.

Targeted searches for CW signals from objects like the Crab and Vela pulsars are very inexpensive. The reason is that since their sky position and frequency evolution are known from electromagnetic observations, one only needs to search for a single waveform or a small number of waveforms around it. About 200 known pulsars rotate at frequencies such that the expected gravitational wave (GW) signal falls into the high sensitivity band of LIGO and Virgo. Searches for GW signals from these objects have been systematically carried out throughout all observing runs [3–6].

At the other end of the spectrum of possible searches are the so-called all-sky searches where one has no prior information on any specific source. Typically these searches span broad ranges of signal frequencies, tens or hundreds of Hz, source positions anywhere in the sky and spindown parameter ranges varying by up to two orders of magnitude [1, 2, 7, 8].

Somewhere in between these two extremes lie the *directed* searches in which we look for signals from an interesting sky region or sky point, e.g. Cas A [9] and generally young supernova remnants [10], the Orion Spur [11], Globular cluster NGC 6544 [12], the low mass x-ray binaries Scorpius X-1 and XTE J1751-305 [13], and the Galactic center [14]. In these searches the computational cost can still be very high because many different waveforms may have to be searched for, corresponding to different signal phase evolutions.

Even for supernova remnants which are very well localized (i.e. for which the localization uncertainty is smaller than the sky resolution of the GW search), the breadth of different waveforms that we would need to search over, corresponding to different frequency and spindown parameters values, is large enough that a single coherent search is computationally unfeasible. We resort then to semi-coherent search schemes [15–21]. In these, the data set is split into N shorter segments of duration $T_{\rm coh}$. Each of these segments is matched with signal templates coherently and in the end the results from these segments are combined incoherently. By reducing the coherent duration $T_{\rm coh}$, at fixed mismatch, the resolution of the waveform parameters becomes coarser and the computing cost is reduced. But this comes at the cost of a reduced sensitivity. In realistic scenarios, we do not have unlimited

computational power available to us. At fixed computing power and for a given data set, there are trade-offs to be made between the length of the coherent segment duration, the resolution in the waveform template bank (the mismatch) and the breadth of the parameter space that we search. Furthermore, these trade-offs should also fold in what we know about the targeted objects, how far they are and how old they are, and the priors that we hold on their deformation and rotation frequency. In computationally limited problems like the ones we are dealing with here, these trade-offs are critical; making the wrong choices could lead us to miss detectable signals and waste significant computational resources.

In [22] we presented an optimization method which, for a given computational budget, a choice of astrophysical priors, and the measured computational efficiency of the search software, tells us how to distribute computing resources for directed CW searches by maximizing the detection probability. The search parameters (including, but not limited to, $T_{\rm coh}$ and N) which will be called *search set-ups*, determine both the computational cost and sensitivity of the search. The optimization method determines the optimal search set-ups for the different targets at different waveform frequencies. The basic idea is to break the large parameter space of the allowed gravitational signals into many small *cells*, and to estimate the computing cost and the detection probability for each cell. Note that the computing cost for each cell depends on parameters related to the search pipeline and software. In particular it depends on the aforementioned coherent segment duration $T_{\rm coh}$, the spacing of the template bank (these are part of the search "set-up"), and also the computational efficiency of the software. We use the detection probability and computational cost to rank what cells and with what search set-up should be searched so that the overall detection probability is maximized. This is done with a linear programming technique. The interplay between the different factors is not trivial and the results are not easy to guess without using the optimization method.

A semi-analytic optimization scheme [23] was previously used to design a directed search for continuous signals from Cas A [9]. Such a scheme minimizes the smallest detectable signal at fixed parameter space and fixed computing cost. In comparison to that scheme, the advantage of the optimization approach [22] that we use here, is that not only does it provide the optimal search set-up, but it also tells us where in parameter space we should spend the computing budget and on what targets. Furthermore the astrophysical priors are explicit, forcing us to spell out what they are. In previous schemes, including [23], the interplay between different priors is often folded-in in a priori choices that are not transparent.

In this paper we present the first application of [22] to set up a directed search by using the first observational run (O1) data of the Advanced LIGO detectors [24]. Following [22], we investigate different astrophysical targets and priors and consider a range of search set-ups for different coherent segment durations, and optimize over all these. We expand with respect to [22] because i) whereas in [22] for each coherent segment duration we considered a single grid spacing combination, here we consider for each coherent segment duration many different combinations of grid spacings and optimize over these, ii) for each grid spacing combination, we fold in the *measured* mismatch distributions obtained from the existing search codes for all the grid spacings and set-ups considered, rather than the analytical estimate which can have very large errors, iii) we revise the computational cost model for our search software to account for recent enhancements in computational efficiency in the search software, iv) we evaluate the loss in detection efficiency incurred if our estimate for the age of a target object is wrong, and finally v) we introduce simplifications with respect to the strictly optimal solution based on practical considerations and estimate the impact on the detection probability.

The plan of this paper is as follows: in Section II we recall the basics of the signal that we want to detect, and introduce quantities that will be referred to in the rest of the paper. The search that we want to perform is introduced in Section III, followed by Section IV where we discuss the astrophysical priors on the signal population. In Section V we carry out the optimization for the three supernova remnants and finally, in Section VI, we summarize and discuss the main results.

II. THE EXPECTED GRAVITATIONAL WAVE SIGNAL

For any plane gravitational wave in standard general relativity, we can choose a wave frame transverse to the direction of propagation such that the two polarizations have the form

$$\begin{aligned} h_+(t) &= A_+ \cos\Phi(t) \\ h_\times(t) &= A_\times \sin\Phi(t). \end{aligned} \tag{1}$$

The phase $\Phi(t)$ of the waves that we target with LIGO is a rapidly varying function of time while the amplitudes $A_{+,\times}$ are generally slowly varying. In fact, consider a rapidly rotating neutron star and let ι be the angle between the total angular momentum of the star and the direction from the star to Earth. The amplitudes of the signals of Eq. (1) are constant over time:

$$\begin{aligned} A_+ &= \frac{1}{2} h_0 (1+\cos^2\iota) \\ A_\times &= h_0 \cos\iota. \end{aligned} \tag{2}$$

Here h_0 is the intrinsic gravitational wave amplitude. The signal at the detector, $h(t)$, is a superposition of the two polarizations

$$h(t) = F_+(\alpha,\delta,\psi;t) h_+(t) + F_\times(\alpha,\delta,\psi;t) h_\times(t), \tag{3}$$

where $F_+(\alpha,\delta,\psi;t)$ and $F_\times(\alpha,\delta,\psi;t)$ are the detector beam pattern functions for the two polarizations. Here

(α, δ) are the right-ascension and declination for the source and ψ is the orientation of the wave-frame with respect to the detector frame. Due to Earth's rotation around its axis and its orbit around the Sun, the relative orientation between the detector and the source is changing continuously which makes $F_{+,\times}$ time varying.

$\Phi(t)$ is the phase of the gravitational wave signal at time t. If $\tau_{\rm SSB}$ is the arrival time of the wave with phase $\Phi(t)$ at the solar system barycenter, then the phase has the form

$$\Phi(\tau_{\rm SSB}) = \Phi_0 + 2\pi[f_0(\tau_{\rm SSB} - \tau_{0\rm SSB}) + \frac{1}{2}\dot{f}_0(\tau_{\rm SSB} - \tau_{0\rm SSB})^2 + \frac{1}{6}\ddot{f}_0(\tau_{\rm SSB} - \tau_{0\rm SSB})^3 + \cdots]. \quad (4)$$

The transformation between detector time $\tau_{\rm De}$ and SSB time $\tau_{\rm SSB}$ is

$$\tau_{\rm SSB}(\tau_{\rm De}) = \tau_{\rm De} + \frac{\mathbf{r}(\tau_{\rm De}) \cdot \mathbf{n}}{c} + \Delta_{\rm E\odot} - \Delta_{\rm S\odot}\,, \quad (5)$$

where $\mathbf{r}(\tau_{\rm De})$ is the position vector of the detector in the SSB frame, $\mathbf{n}$ is the unit vector pointing to the source, and c is the speed of light; $\Delta_{\rm E\odot}$ and $\Delta_{\rm S\odot}$ are respectively the relativistic Einstein and Shapiro time delays.

III. THE SEARCH

Based on the results of [22], we concentrate on three targets, Vela Jr., Cas A and G347.3, and illustrate the design of three directed searches for continuous GW signals from these. These three targets are supernova remnants and are believed to harbour a neutron star. The electromagnetic observations identify these as point sources. The position of Vela Jr. comes from Chandra X-ray satellite data [25]. The estimates of its age and distance are uncertain [26]. Cassiopeia A (Cas A in short) is one of the youngest known supernova remnants. The position of Cas A comes from Chandra data [27], the distance from [28] and the age from [29]. G347.3 is a target which is not only close (1.3 kpc [30]) but also young (1600 years old [31]). The position of G347.3 is also from Chandra data [32].

We assume using the volunteer computing project Einstein@Home [33], data with the average noise of the Advanced LIGO detectors during their first observational run (designated as O1), and with the duration and average duty factor of the LIGO O1 data.

A. Search method

As mentioned earlier, we take a semi-coherent approach in which the data is divided into shorter segments each spanning the same observation time $T_{\rm coh}$. Each of these segments are analyzed coherently and afterwards combined incoherently. The coherent analysis in each segment uses the $\mathcal{F}$ statistic introduced in [34, 35]. The final semi-coherent detection statistic is the average of the $\mathcal{F}$-statistic for each segment which shall be denoted as $\widehat{2\mathcal{F}}$ as implemented in the GCT (Global correlation transform) method [20], used in the most recent Einstein@Home papers [1, 2, 9].

B. Template banks and mismatch distributions

As mentioned earlier, each waveform is defined by the parameters $(h_0, f, \dot{f}, \ddot{f}, \alpha, \delta, \psi, \iota, \Phi_0)$. However the coherent search method analytically maximizes over the parameters $(h_0, \psi, \iota, \Phi_0)$ and so we only need to explicitly search over $(f, \dot{f}, \ddot{f}, \alpha, \delta)$. Since we consider directed searches the sky position is known and the different waveforms are determined by only $(f, \dot{f}, \ddot{f})$.

With the term "template bank" we indicate the collection of waveforms that we explicitly search for. These are defined by values of the waveform parameters $(f, \dot{f}, \ddot{f})$. The spacings in these parameters in principle has to be fine enough that a real signal, with waveform parameters lying between adjacent points in the template bank would still be detectable.

Two grids are set-up: a coarse grid, used for the coherent searches, and a fine grid, used for the incoherent searches. The frequency spacing δf is the same for both the coherent and incoherent searches whereas the frequency derivative spacings are refined by a factor of $\gamma^{(1)}$ and $\gamma^{(2)}$ going from the coherent to the incoherent grids: $\delta\dot{f}_f = \frac{\delta\dot{f}_c}{\gamma^{(1)}}$ and $\delta\ddot{f}_f = \frac{\delta\ddot{f}_c}{\gamma^{(2)}}$. The total number of waveforms that are searched is $N_{\rm fine} = \gamma^{(1)}\gamma^{(2)}N_{\rm coh}$ with $N_{\rm coh} = N_{f_c}N_{\dot{f}_c}N_{\ddot{f}_c}$, where N_{f_c}, $N_{\dot{f}_c}$ and $N_{\ddot{f}_c}$ are the number of coarse grid points in f, $\dot{f}$ and $\ddot{f}$ respectively.

The grid spacings $\delta f_c, \delta\dot{f}_c, \delta\ddot{f}_c$ are parametrized by the parameters m_f, $m_{\dot{f}}$, $m_{\ddot{f}}$ as follows [36]:

$$\delta f_c = \frac{\sqrt{12 m_f}}{\pi T_{\rm coh}}\,, \quad (6)$$

$$\delta\dot{f}_c = \frac{\sqrt{180 m_{\dot{f}}}}{\pi {T_{\rm coh}}^2}\,, \quad (7)$$

$$\delta\ddot{f}_c = \frac{\sqrt{25200 m_{\ddot{f}}}}{\pi {T_{\rm coh}}^3}\,. \quad (8)$$

Every template bank for a given search set-up ($T_{\rm coh}$ and N) is characterized by its mismatch distribution. The mismatch μ is the quantity that measures how much signal-to-noise may be lost due to the mismatch between a signal parameters and the discrete template bank. The finer the template bank is, in general the smaller the mismatch. The mismatch distribution can be measured by simulating signals (no noise) and measuring the signal-to-noise ratio $\rho^2_{\text{no-mismatch}}$ associated with a search performed with a perfectly matched template (i.e. a template with parameters identical to the signal parameters) and the signal-to-noise ratio $\rho^2_{\rm mismatch}$ associated with the

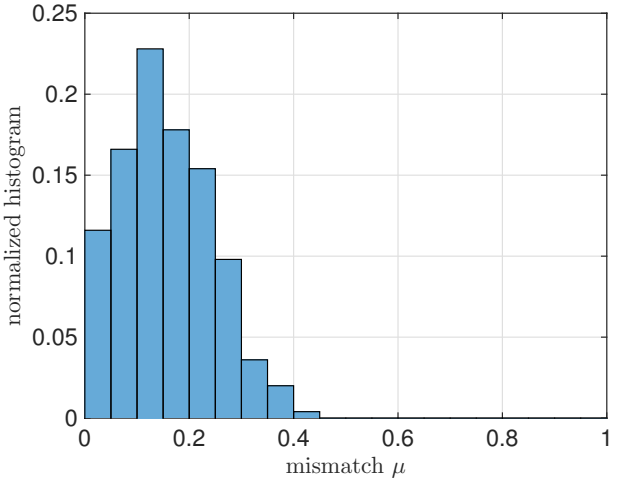


FIG. 1: This histogram shows the mismatch distribution for the grid spacing: $m_f = 0.15$, $m_{\dot{f}} = 0.3$, $m_{\ddot{f}} = 0.003$, $\gamma^{(1)} = 8$, $\gamma^{(2)} = 20$ and $T_{\rm coh}$=15 days for Vela Jr and it is the set-up that was eventually chosen for the search. The average measured mismatch is $\langle\mu\rangle = 15.8\%$.

maximum of the detection statistic from a search performed with the template grid that we want to characterize:

$$\mu := \frac{\rho^2_{\text{no-mismatch}} - \rho^2_{\text{mismatch}}}{\rho^2_{\text{no-mismatch}}}. \tag{9}$$

The signal-to-noise ratio ρ^2 for a semi-coherent $\mathcal{F}$-statistic search with N segments is connected to the expectation value of $\hat{\mathcal{F}}$ as follows: $E[2\hat{\mathcal{F}}] = \rho^2 + E[n]$, where $E[n]$ is the expected value of noise alone.

We derive the mismatch distribution for a given search set-up $T_{\rm coh}$, N and a template bank defined by $m_f, m_{\dot{f}}, m_{\ddot{f}}, \gamma^{(1)}$ and $\gamma^{(2)}$ by injection-and-recovery Monte Carlos following Eq. (9). In particular, 500 artificial signals are produced with gaps simulating the realistic output from the LIGO O1 detectors. These data streams are jointly searched with our standard semi-coherent search[1]. Because these injection-and-recovery Monte Carlos are performed on the noise-free data streams, the computed values of the search code correspond to the expectation values of the statistic, $E[\hat{\mathcal{F}}]$, and therefore we obtain the mismatches directly from Eq. (9). The range of the parameters of the 500 signals is given in Table I. We have considered over 2000 different spacings. The ranges for these spacings are listed in Table II. For illustration purposes, Fig. 1 shows the mismatch distribution for $T_{\rm coh} = 15$ days, $m_f = 0.15$, $m_{\dot{f}} = 0.3$, $m_{\ddot{f}} = 0.003$, $\gamma^{(1)} = 8$ and $\gamma^{(2)} = 20$. The average measured mismatch is $\langle\mu\rangle = 15.8\%$ and this is a typical value for a deep broad parameter space CW survey (cfr. for instance with the average mismatch of the last stage follow-up in [1] which lies at ~ 0.13).

C. Search software timing

In computationally limited problems, one needs a rational process for allocating computational resources amongst different competing proposals. For example, should we allocate equal resources for each of Vela Jr., Cas A and G347.3? If not, then what is the optimal distribution of computing power?

An important ingredient in the optimization is an accurate estimation of the run-time of the search pipeline for a given search set-up over a given parameter space. CW searches are the most computationally expensive gravitational wave searches and a great effort has been employed to optimize them. As a result, a well developed timing model has been developed for our main search pipeline. Furthermore, the use of Einstein@Home demands that we are able to predict the run-time of the work-units assigned to each of the host machines, which adds a further incentive to characterize the software accurately.

The time $\tau_{\rm total}$ that it takes to perform a search in a parameter-space volume covered by $N_{\rm coh}$ coarse grid templates and $N_{\rm fine}$ fine grid templates, using data from $N_{\rm det}$ detectors, divided among N segments can be written as:

$$\begin{aligned}\tau_{\rm total} =& N N_{\rm det} N_{\rm coh} \tau_{\rm RS} + N N_{\rm inc} \tau_{\rm sum\mathcal{F}} \\ &+ N_{\rm inc} \tau_{\rm Bayes} + N_{\rm can} \tau_{\rm Recalc}.\end{aligned} \tag{10}$$

The timing coefficients $\tau_{\rm RS}$, $\tau_{\rm sum\mathcal{F}}$, $\tau_{\rm Bayes}$ and $\tau_{\rm Recalc}$ are determined based on computing time measurements executed with various search set-ups on `Intel Xeon E3-1231 v3` CPUs [2].

The first term in Eq. (10) is the cost of the coherent step. When searching over several thousand signal frequency bins corresponding to the same sky position and spindown values, the best algorithmic implementation of the $\mathcal{F}$-statistic is obtained by carrying out the frequency demodulation of the signal by resampling a downsampled time-series according to $\tau_{\rm SSB}$ and then performing an FFT [34, 38]. Our most recent enhancement of the search codes uses this resampling + FFT method. $\tau_{\rm RS}$ is the time that it takes to calculate a value of the coherent detection statistic, corresponding to a single template, using data from a single detector and from a single segment. It can be written as [39] :

$$\tau_{\rm RS} = \tau_{\rm Fbin} + \frac{N^{\rm FFT}_{\rm samp}}{N_{\rm Fbin}}(\tau_{\rm FFT} + R\tau_{\rm spin}), \tag{11}$$

[1] The artificial data is created with `lalapps_Makefakedata_v4`. The search is performed using `lalapps_HierarchSearchGCT` (GCT). Both programs are part of the LIGO Algorithm Library (LALSuite) [37].

[2] http://ark.intel.com/de/products/80910/Intel-Xeon-Processor-E3-1231-v3-8M-Cache-3_40-GH

TABLE I: Injection parameters used in mismatch investigation

Parameter	Range
Frequency [Hz]	$151\,\mathrm{Hz} \leq f \leq 152\,\mathrm{Hz}$
First spin-down [Hz/s]	$-10^{-7}\,\mathrm{Hz/s} \leq \dot{f} \leq -10^{-12}\,\mathrm{Hz/s}$
Second spin-down [Hz/s^2]	$10^{-22}\,\mathrm{Hz/s^2} \leq \ddot{f} \leq 7\times10^{-17}\,\mathrm{Hz/s^2}$

Notes: Parameters of the fake signals used to derive the mismatch distributions. f is uniformly randomly distributed; $\dot{f}$ and $\ddot{f}$ are log-uniformly randomly distributed.

TABLE II: Template bank parameters

Parameter	Range
T_{coh} [days]	10, 15, 20, 30, 60
m_f	$0.1 \leq m_f \leq 1.0$
$m_{\dot{f}}$	$0.1 \leq m_{\dot{f}} \leq 1.0$
$m_{\ddot{f}}$	$0.001 \leq m_{\ddot{f}} \leq 1.5$
$\gamma^{(1)}$	$1 \leq \gamma^{(1)} \leq 50$
$\gamma^{(2)}$	$5 \leq \gamma^{(2)} \leq 100$

Notes: The total observation duration T_{obs} is 120 days. It is divided into $N \approx T_{\mathrm{obs}}/T_{\mathrm{coh}}$ coherent segments, the "≈" due to there being gaps in the science-quality data.

where τ_{Fbin} is the time spent on operations on each output frequency bin, τ_{FFT} is the time spent on the FFT per sample $N^{\mathrm{FFT}}_{\mathrm{samp}}$ of the resampled and zero-padded time series, and τ_{spin} is the time spent per sample of the SSB-frame resampled time series without zero-padding, with length $R\,N^{\mathrm{FFT}}_{\mathrm{samp}}$ where $R \leq 1$. Zero-padding is used in order to obtain the desired frequency resolution of the resulting $\mathcal{F}$-statistic. The timing coefficients of Eq. (11) are: $\tau_{\mathrm{Fbin}} = 6.0\times10^{-8}$ s, $\tau_{\mathrm{FFT}} = 3.3\times10^{-8}$ s, $\tau_{\mathrm{spin}} = 7.5\times10^{-8}$ s.

The second term in Eq. (10) is the cost of the incoherent step: For every fine grid point we sum N detection statistic values. For $2 \leq N \leq 12$ the timing coefficient $\tau_{\mathrm{sum}\mathcal{F}}$ is

$$\tau_{\mathrm{sum}\mathcal{F}} = 7.28\times10^{-9} - 3.72\times10^{-10}N\ \mathrm{(s)}\ . \qquad (12)$$

Since the efficiency of adding up detection statistic values increases with the number of segments, there is a negative term in Eq. (12) proportional to number of segments N. Note that Eq. (12) has been measured by timing the search on set-ups where N varies between 2 and 12, which is the range of interest for our data set, hence, it may not hold for N values outside of this range. Eq. (12) has been obtained from the linear fitting of 23 timing trials with a norm of residuals 2.4×10^{-9}.

As done in [1, 2] the main detection statistic is augmented with variants that are robust to detector artefacts, namely the line-robust $\hat{B}_{\mathrm{S/GL}}$ and the transient-robust $\hat{B}_{\mathrm{S/GLtL}}$, as well as a detection statistic which is sensitive to some types of transient signals, $\hat{B}_{\mathrm{tS/GLtL}}$ [40–42]. The third term in Eq. (10) is the time to compute these specialized statistics *given* the single detector and multi detector coherent detection statistic values (see Eq. (13) of [42]). This is the reason why τ_{Bayes} is independent of the number of segments. $\tau_{\mathrm{Bayes}} = 4.4\times10^{-8}$ s is also obtained from the 23 timing trials. For set-ups with just a few segments, the cost of computing these various statistics can be larger than the cost of the incoherent step.

The last term in Eq. (10) is the computing cost for the recalculation all these detection statistics at the exact fine grid point in all the coherent segments. This is done only for the N_{can} candidates that are in the top list[3]. In the last few Einstein@Home searches the number of candidates in the top-list is $N_{\mathrm{can}} = O(1000)$ and this recalculation cost is negligible with respect to the costs of the other three terms.

The computing time, and consequently the values of the timing coefficients, in general depend on the CPU on which the search is performed. Since the volunteer computing project Einstein@Home comprises a broad range of different CPUs, as we optimize this search for running on Einstein@Home for a predetermined length of time, we need to determine how much computing power that corresponds to. A timing analysis based on duration of the work units (WUs) of the O1 all-sky low frequency Einstein@Home search [8], yields the results shown in Fig. 2: run times are bi-modally distributed, with a mode centered at 8-10 hours, the other at 24-26 hours. Based on this we divide the host population of Einstein@Home into two types: hosts that showed a runtime of less than 14 hours were put in one category (A) and hosts that needed more time than that were placed in the other category (B). The CPU models in host class (A) have an average 8-hours runtime which is equivalent to the fast nodes on the ATLAS computational cluster at the Albert Einstein Institute in Hannover [43]. Therefore the

[3] The top list is the list of top candidates that is returned by the Einstein@Home volunteer computer to the main Einstein@Home server. Typically multiple top lists will be returned, each ranked according to a different detection statistic.

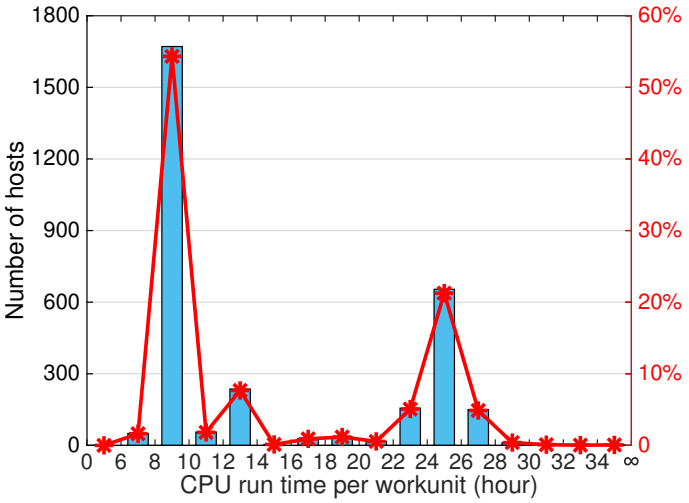


FIG. 2: Distribution of the run-times of the Einstein@Home O1 low frequency all-sky search work units. The data in this figure is taken from a sample of 3076 hosts out of total 28764 hosts of Einstein@Home.

timings from fast nodes on ATLAS can be ideally used for these Einstein@Home (A)-cores. Based on the O1 low frequency all-sky search we estimate there are about 5300 cores in category (A) and a similar number of cores in category (B). In the following we will assume using only the (A)-type hosts of Einstein@Home, and hence that one Einstein@Home month (EM) corresponds to 5300 (A)-type cores used continuously for 1 month.

IV. ASTROPHYSICAL PRIORS

An important feature of our optimization scheme is that it forces us to explicitly incorporate our astrophysical priors on the signal parameters. In particular, we need to choose the astrophysical priors on the following parameters for each target:

- the age and distance of the target sources. The age influences the spindown range that should in principle be searched. The distance of the source directly influences the amplitude h_0 of the signal. For some targets such as Cas A there is very little uncertainty in the distance and/or age of the object, so the prior is chosen as a delta function. Other sources like Vela Jr., for instance, have large uncertainty in age and distance. However as was done in [22], we pick delta-functions at the extremes of the possible range. In doing this we might also include non-physically motivated age-distance combinations. We do this only to give a sense of the impact of different priors on the final results.

- The star's ellipticity ε and the fraction x of spindown energy carried away in GWs influence the GW signal amplitude h_0. For the former we pick a log-uniform distribution; for the latter we pick a value inspired by the upper limits measured with targeted searches [3]

- The P_c depends not only on h_0, but also on the prior probability $P(f, \dot{f}, \ddot{f})$ that the actual signal is in the cell defined by specific values of $(f, \dot{f}, \ddot{f})$.

The results in [22] indicate that the distance of an object is the most important parameter in determining its GW detectability. Among the targets considered in [22] targets Vela Jr. is the only source which has large uncertainties both in age and distance. However, even assuming a fairly pessimistic value of 750 pc for the distance of Vela Jr. [44], it still contributes the most to the total detection probability with respect to all other targets. We pick the four extremes for the priors on age and distance of Vela Jr.: close and young (CY), close and old (CO), far and young (FY) and far and old (FO) Vela Jr.. It is important to keep in mind that the astrophysically viable alternatives are CY and FO and there is no support for CO and FY [4]. However, we shall include all four alternatives for the purpose of illustrating the impact on the optimization scheme. Realistic searches will obviously consider only CY and FO.

Table III details the parameters of the three objects that we consider in this paper. Searches for signals from one of the youngest known SNR, Cas A, have been carried out with LIGO data [9, 10, 45]. The reason for targeting this source is that a young object is more likely to be spinning down faster and hence there is more kinetic energy that could potentially be radiated away in GWs. Here we will see that Cas A is the third source that contributes the overall detection probability. Vela Jr. and G347.3 contribute much more to the total detection probability and are thus more promising targets. While [10] presents searches and upper limits for Vela Jr. and G347.3, these have not been the primary targets for any CW search to date. Thus there have not been any deep CW searches for these two objects so far.

The ellipticity ε is the least known parameter so here we take a flat probability density on $\log \varepsilon$ within a conservative range of values. We target weak signals and hence the maximum value of ε we allow in the ith cell is:

$$\varepsilon_i^{\max} = \min(\, \varepsilon_i^{\rm sd},\ \varepsilon_i^{\rm age},\ 10^{-6}) \,, \tag{13}$$

where $\varepsilon_i^{\rm sd}$ and $\varepsilon_i^{\rm age}$ are the spin-down ellipticity upper limit and the spindown age-based ellipticity upper limit

[4] To estimate the age for Vela Jr., one measures its angular size θ and expansion rate R. It follows that the ratio of the distance D and the age τ is given by $D/\tau = R/\theta$. Thus one does not measure D or τ independently but only the ratio and so a larger D implies a larger τ. Therefore we have astrophysical support for CY and FO and not for CO or FY.

TABLE III: Point source targets considered in this paper

SNR G name	Other name	Point source J	$D_{\rm kpc}$	$\tau_{\rm kyr}$
111.7−2.1	Cas A	232327.9+584842	3.3–3.7	0.31–0.35
266.2−1.2	Vela Jr.	085201.4−461753	0.2–0.75	0.7–4.3
347.3−0.5		171328.3−394953	1.3	1.6

respectively:

$$\varepsilon^{\rm sd} = \sqrt{\frac{5c^5}{32\pi^4 G}\frac{x|\dot{f}|}{If^5}}\,. \tag{14}$$

and

$$\varepsilon^{\rm age} = \frac{c^2}{16\pi^2 f^2}\sqrt{\frac{10c}{GI\tau}}\,. \tag{15}$$

In Eqs. (14) and (15), f is the instantaneous frequency of the emitted GW signal, G Newton's constant, c the speed of light, and τ the age of the source. I is the the moment of inertia and we use its standard value $10^{38}\,{\rm kg\,m^2}$ for all the results in this paper. x is the fraction of spin-down energy loss due to GW emission. The latest observational limits on GW emission from the Crab and Vela pulsars constrain x to less than 0.2% and 1% respectively [3]. In this paper, we will assume $x = 1\%$. There is no guarantee that the Vela and Crab results apply to these other objects, but right now they are the only measurements at hand. According to the results of [46], the realistic maximum value of ε is expected to be smaller than 4×10^{-6}. Hence the third limit we use in this paper is 10^{-6}. We take $\varepsilon^{\rm min} = 10^{-14}$ because deformations of a compact star due to the internal magnetic field (at least 10^{11} G) are not expected to be smaller than $\sim 10^{-14}$ [47]. Based on the above discussion, our prior $p(\varepsilon)$ is:

$$p(\varepsilon) = \begin{cases} \frac{1}{\varepsilon}\frac{1}{\log(\varepsilon^{\rm max}/\varepsilon^{\rm min})} & \varepsilon^{\rm min} < \varepsilon < \varepsilon^{\rm max} \\ 0 & \text{elsewhere}\,. \end{cases} \tag{16}$$

Since the GW frequencies emission frequencies are unknown, our search encompasses a large range; namely, from 20 to 1500 Hz. For a given f, the $\dot{f}$ and $\ddot{f}$ ranges are determined by the fiducial age of the source τ:

$$\begin{cases} 20\,{\rm Hz} \le f \le 1500\,{\rm Hz} \\ -f/(n-1)\,\tau \le \dot{f} \le 0\,{\rm Hz/s} \\ 0\,{\rm Hz/s^2} \le \ddot{f} \le nf/\tau^2. \end{cases} \tag{17}$$

n is the braking index. If the frequency evolution follows $\dot{f} \propto f^n$, the second order spindown is then $\ddot{f} = n\dot{f}^2/f$. In the second equation we take $n = 2$ to encompass the broadest range of $\dot{f}$ values. In the third equation we assume a braking index $n = 5$, corresponding to phase evolution purely due to GW emission, and a constant $\dot{f}$ value of $-f/\tau$. For all other mechanisms $n < 5$ and in particular for pure dipole electromagnetic emission $n = 3$ (see e.g. [48]). Therefore, our search ranges for $\dot{f}$ and $\ddot{f}$ in (17) encompass all combinations of emission mechanisms. Since none of the quantities that determine the detection probability at a given frequency, such as the maximum ellipticity or the noise level, depend on $\ddot{f}$ we drop the $\ddot{f}$ dependence in the signal probability density $P(f,\dot{f})$, equivalent to a uniform prior on the $\ddot{f}$ range. We consider both uniform and log-uniform priors on the search range of f and $\dot{f}$ reflecting our ignorance on those signal parameters.

In this search, we do not search over the third order spin-down $\dddot{f}$ because, even for the youngest target, a search over third order spin downs is not necessary for the coherent and observation times that we are considering here.

V. THE OPTIMIZATION

The optimization scheme is introduced in [22]. The starting point is to divide the parameter space into non-overlapping small cells such that the computing cost to search each cell and the resulting detection probability are roughly constant within each cell. The cost for each should also be much smaller than the full computing cost budget available to us. For each cell, we calculate the computing cost and detection probability. For each cell we define the efficiency, that is the ratio of detection probability to computing cost. In the absence of any constraint apart from the total budget, one could proceed simply by picking the most efficient cells till the computing budget is exhausted. However, we do have an additional constraint, namely that we do not want to search the same parameter space cell with multiple search set-ups. It is shown in [22] how this constraint can be included in the optimization using linear programming techniques.

The search set-ups are defined by 6 parameters: the segment coherent duration $T_{\rm coh}$, the nominal mismatch parameters m_f, $m_{\dot{f}}$, $m_{\ddot{f}}$, and the refinement factors $\gamma^{(1)}$ and $\gamma^{(2)}$ (see details in Section III B). For each given search set-up we can compute the mismatch distribution, i.e. the distribution of the fractional loss in signal-to-noise ratio due to the discreteness of the template bank[5].

[5] Mismatch distributions of these set-ups are computed for Vela

The optimization not only determines the best search set-ups but also how to distribute the computing budget among the different astrophysical targets and over the parameter space. In [22] we presented this optimization scheme under a set of simplifications. We now take [22] as the starting point for an actual search design but here we forego the simplifications that we made in [22] and take into account issues of feasibility and practicality. We shall take rectangular cells in $(f, \dot{f})$ space, and each cell will be 10 Hz wide in frequency and 10^{-9} Hz/s in $\dot{f}$. To completely specify the parameter space cells, we need to include $\ddot{f}$, and each of the $(f, \dot{f})$ cells are allowed to take all permissible values of $\ddot{f}$. As mentioned above, for each cell, each search set up, and every choice of astrophysical prior, we calculate the computing cost and detection probability. Since it is not possible to predict the mismatch distribution of the GCT search for a given set-up, these are produced from thousands of injection-and-recovery Monte Carlos.

In principle, the above ingredients are sufficient for the optimization scheme. However, certain obvious simplifications can be made based on practical considerations:

- Among all the set-ups with the same $T_{\rm coh}$, for every value of the computing cost, we select the set-up which has the lowest average mismatch. In our case, as we shall describe in greater detail below, this results in 71 seeded set-ups for all the different values of $T_{\rm coh}$ considered.

- Among all the cells, each with 71 set-ups, we select those so that the sum of relative detection probabilities constrained by the available computing budget, is maximized. This is done through linear programming.

- Since we find that the optimal choice of cells spans a very broad frequency range, we determine the loss in detection probability caused by extending the search parameter space to include the broadest frequency band. We further evaluate the loss incurred with respect to the optimal solution due to utilizing the same set-up across all cells for every astrophysical target. Both these choices, when viable, greatly simplify the post-processing of the results.

- We compare the final results from different sets of priors, estimate the loss in detection efficiency due to having optimized assuming a wrong age and choose the set-up such that this loss is the smallest.

In the next sections we describe the above steps in greater detail.

Jr., since it is responsible for most of the detection probability

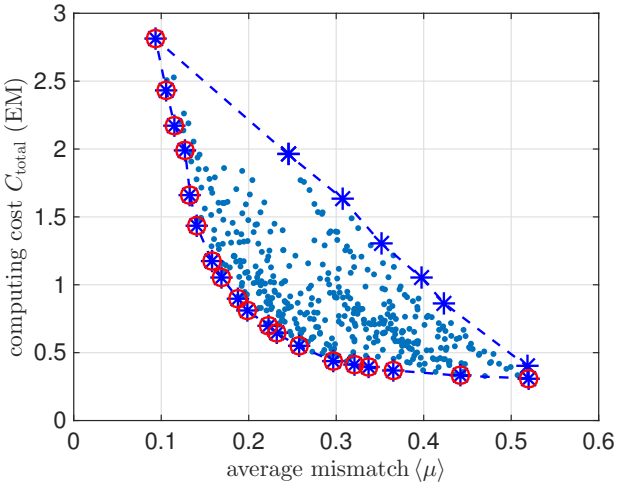


FIG. 3: 432 set-ups with $T_{\rm coh} = 15$ days. The 19 blue stars in red circles are selected for further consideration as explained in the text.

A. Set-ups: primary selection

Running the optimization scheme on ~ 2000 set-ups is computationally too burdensome. We hence down-select, among the ones with the same $T_{\rm coh}$, those that yield the lowest average mismatch at fixed computing cost over the entire prior parameter range of Vela Jr.. In principle this selection should be done separately for every target, but here we simplify the procedure in this manner because Vela Jr. contributes the bulk of the detection probability.

Fig. 3 shows computing cost and average mismatch for 432 set-ups corresponding to different grid spacings and $T_{\rm coh} = 15$ days. Out of these, 19 are selected which have the lowest measured average mismatch at fixed computing cost. Considering different values for $T_{\rm coh}$ we select 71 set-ups : 13 are from $T_{\rm coh} = 10$ days, 19 from $T_{\rm coh} = 15$ days, 15 from $T_{\rm coh} = 20$ days, 14 from $T_{\rm coh} = 30$ days and 10 from $T_{\rm coh} = 60$ days. Since the core of the optimisation procedure is linear programming and its computing time grows at least polynomially with the number of tested set-ups [49], it is easy to see that going from 2000 set-ups to 71 reduces the cost by at least a factor of $>$ 100. For reference we note that it took $\sim$10 CPU hours to perform the optimisation with the 71 set-ups.

The *measured* average mismatch $\langle\mu\rangle$ of each of these distributions is used to reduce the expected signal-to-noise ratio ρ^2 of a putative signal. More specifically, with respect to Eq. (16) of [22] that used only ρ^2, we now use a more realistic estimate of the actual results of a search:

$$\rho^2 \to (1 - \langle\mu\rangle) \times \rho^2. \tag{18}$$

This estimate folds-in the sensitivity-loss effect of using finite grids and does it realistically because it is based on the measured mismatch distributions of the actual search codes.

B. Optimization under different assumptions

Following Section V of [22], we now use linear programming to determine how to best pick targets, waveform parameter space to search and search set-ups under a set of different assumptions on the age and distance of the source. We consider 71 different set-ups. A significant difference with respect to [22] is that here we consider a number of different grid spacings for the same $T_{\rm coh}$ and, as explained above, *the actual* mismatch associated with each of them.

Fig. 4 shows the results of the optimization under the assumptions that Vela Jr. is at a distance of 200 pc (C), that the signal frequency f and spindown $\dot{f}$ are uniformly distributed within their ranges (defined in Eq. (17)) and with a 3 EM computing budget.

Fig. 4 shows the result of the optimization procedure. Different colors represent search set-ups having different $T_{\rm coh}$. The total detection probability $\mathcal{R}$ is defined in Eq. (49) of [22] as the sum of detection probabilities of the selected parameter space cells. C indicates the total computational cost and in general this will be equal to or smaller than the maximum budgeted 3EM. The results for Vela Jr. for the CY and CO cases can be summarized as follows:

- When Vela Jr.'s age is assumed to be 700 years, the spin-down parameter space to search is large and the $T_{\rm coh}$ chosen set-ups for Vela Jr. are 10 days, 15 days and 20 days (see Fig. 4(a)). Fig. 4(c) shows the Vela Jr. plane of Fig. 4(a) and we can appreciate that more than 10 different set-ups constitute the optimal search for young Vela Jr.

- When Vela Jr.'s age is assumed at 4300 years, the spin-down parameter space shrinks and the $T_{\rm coh}$ chosen set-ups for Vela Jr. are all 20 days (see Fig. 4(b)). This is the CO case and we emphasize again that there is no astrophysical support for this scenario. We include this to illustrate the effect of the priors on the optimization results.

The total detection probability $\mathcal{R}$ corresponding to the two different age priors is also different: 18.7% and 16.1% for the young and for the old Vela Jr., respectively. Under either assumptions Vela Jr. contributes the bulk of the total probability: 14.4% out of 18.7% and 11.5% out of 16.1%. However in the former case the Vela Jr. search uses up 38% of the computing budget whereas in the latter it only uses 7%.

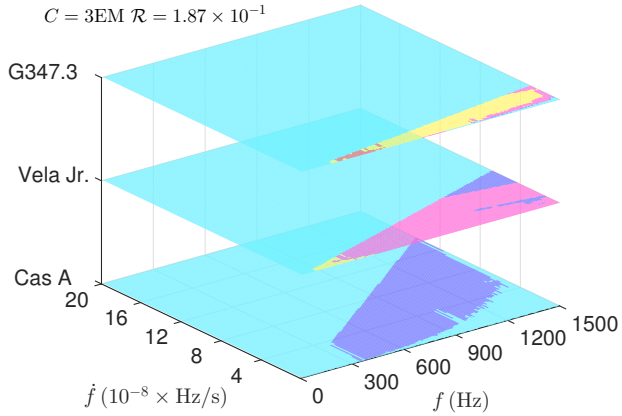
Figs. 5 reveals further details of the chosen set-ups. In particular panels (a), (c) and (e) show how many cells are searched with each of the different set-ups for G347.3, Vela Jr. (CY) and Cas A, respectively. Panels (b), (d), (f) show the same quantities but under the CO assumption for Vela Jr.. Assuming Vela Jr. is 700 yr old (CY), 20-day set-ups are mostly selected for G347.3, 15-day set-ups mostly for Vela Jr., and only 10-day set-ups for Cas A. If we assume that Vela Jr. is 4300 yr old, the parameter space of Vela Jr. shrinks due to the age limit. Only 20-day set-ups are selected for Vela Jr.. The computing savings incurred due to the smaller parameter space are re-invested in longer $T_{\rm coh}$. The dominant set-ups for G347.3 use shorter $T_{\rm coh}$ compared to the ones derived under the Vela Jr. CY prior. This is due to the fact that, with a smaller parameter space for Vela Jr., the most probability is harvested by exploiting it the most, with long (expensive) $T_{\rm coh}$ and this is balanced by spending less on the other targets. Finally, because of the distance of Cas A is much larger than Vela Jr. and G347.3, the optimal way to distribute the computing budget is by searching Cas A with relatively cheap (less-sensitive) 10-day set-ups.

Since we do not know the age of Vela Jr., the set-up corresponding to which of the two priors should we pick? To answer this question, we investigate the consequences of having picked the wrong prior, namely the impact on the detection probability if we assume that Vela Jr. is 700 yrs old when Vela Jr. is 4300 yr old and if we assume that Vela Jr. is 4300 yrs old when Vela Jr. is 700 yr old. In the first case we search a broader spin-down range than we need to, in order to gather all the detection probability. To do this we use a less-sensitive set-up (shorter $T_{\rm coh}$) for most of the frequency band for Vela Jr. and, partly also for Cas A and G347.3 (see Fig. 5(a), (c), (e)). The consequence of this is that we waste computing power in the high $\dot{f}$ region and lose detection probability due to the shorter $T_{\rm coh}$ set-ups used in the low $\dot{f}$ region. In the second case, we just search the low $\dot{f}$ region with longer Tcoh set-ups (see Fig. 5(d)). We gain some detection probability due to adopting longer Tcoh set-ups in the low $\dot{f}$ of Vela Jr. and in parameter space of Cas A and G347.3. Meanwhile, we lose detection probability because we give up the whole high $\dot{f}$ range of Vela Jr..

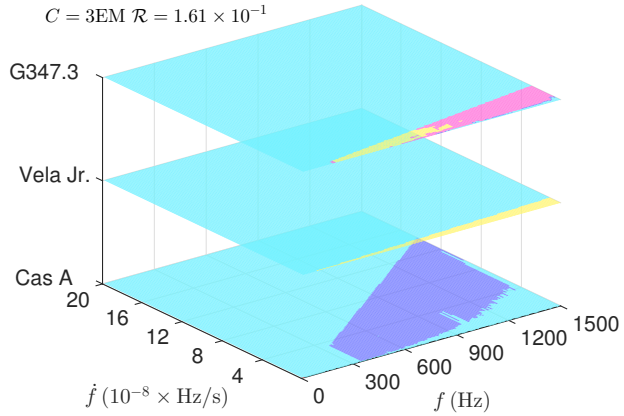
The results are summarized in Table IV. From these it is clear that using the youngest-age prior for Vela Jr. leads to the smallest loss in detection probability if this assumption is wrong, hence we use this prior in our optimizations.

C. The total computing budget

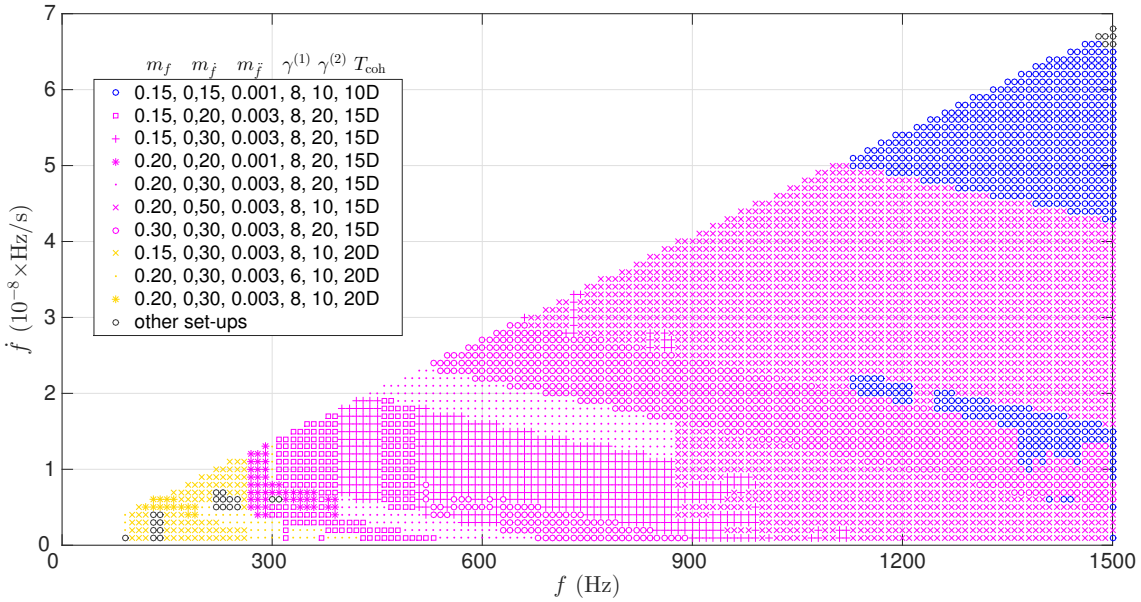
In the previous discussion we found that, if the likelihood of Vela Jr.'s frequency and spindown is uniform between 20 and 1500 Hz, and if Vela Jr. is 700 yrs old and at a distance of 200 pc, by optimally choosing set-ups and target parameter space to search with 3 EMs, the total detection probability $\mathcal{R}$ is 18.7%. It is then natural to ask whether by investing more computing resources we could achieve an even higher detection probability. Assuming a set of computing budgets from 0.1 EM to 12 EM, the optimization procedure yields the $\mathcal{R}$s shown in Fig. 6 as a function of C. Although $\mathcal{R}$ always grows as C increases, the growth rate decreases as $\mathcal{R}$ increases. We identify three stages. In the first stage, when C is from



(a) Assuming Vela Jr. is close and young



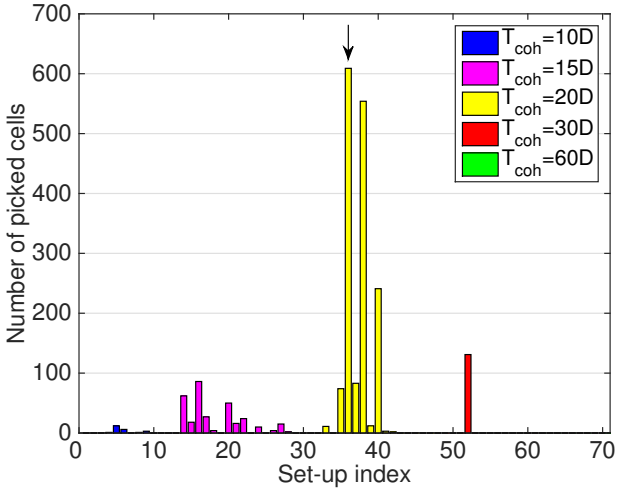
(b) Assuming Vela Jr. is close and old



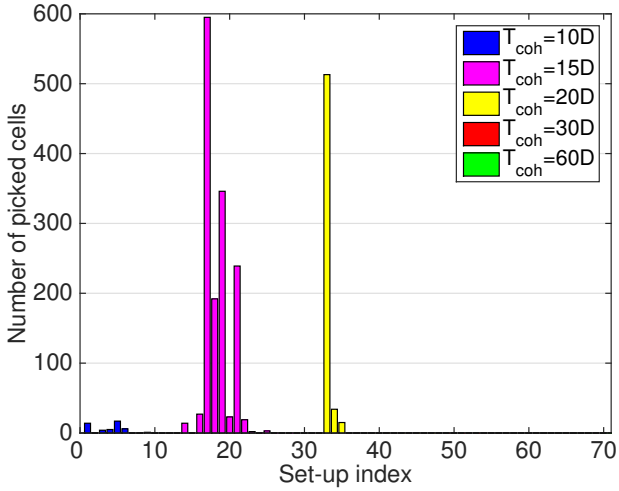
(c) Set-up details on the plane of Vela Jr. (assuming Vela Jr. is close and young)

FIG. 4: Parameter space coverage assuming uniform $f, \dot{f}$ priors, Vela Jr. at 200 pc (C) and 3 EM computing budget. In panel (a) we assume Vela Jr.'s age is 700 years (Y) and in panel (b) we assume that Vela Jr. is 4300 years old (O). The plane is indicated in aqua. Cells in blue are searched with 10-day T_{coh} set-ups, magenta indicates 15-day T_{coh}, yellow the 20-day T_{coh}, red the 30-day T_{coh}, and green the 60-day T_{coh} (although not used in either cases). The computing power used on Cas A, Vela Jr. and G347.3 is 1.36 EM, 1.14 EM and 0.50 EM, respectively, for the set-ups of panel (a), and 1.97 EM, 0.21 EM and 0.82 EM for the set-ups of panel (b). The contribution to the total detection probability from Cas A, Vela Jr. and G347.3 is 1.3%, 14.4% and 3.1% for panel (a), and 1.4%, 11.5% and 3.1% for panel (b). Note that each color represents set-ups which have the same coherent duration, but might differ in the grid spacings. Figure (c) shows the set-up details on Vela Jr. plane in Figure (a). For example, the blue circles represent the cells that need to be searched by using the set-up: $m_f = 0.15$, $m_{\dot{f}} = 0.15$, $m_{\ddot{f}} = 0.001$, $\gamma^{(1)} = 8$, $\gamma^{(2)} = 10$, and $T_{\text{coh}} = 10$ day.

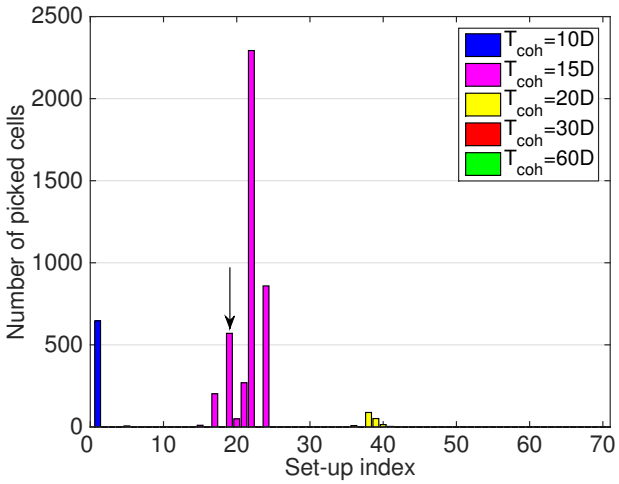
0.1 to 0.4 EM, $\mathcal{R}$ increases very fast. In the second stage when C is from 0.4 to ∼3 EM, $\mathcal{R}$ still increases but not as fast as before. In the last stage when C is larger than ∼3 EM, $\mathcal{R}$ increases even more slowly. In this regime a gain in the $\mathcal{R}$ due to 9 additional EMs, is less than 1%. Based on this we decide to invest around 3 EM in this search, covering the first two stages of $C(\mathcal{R})$.



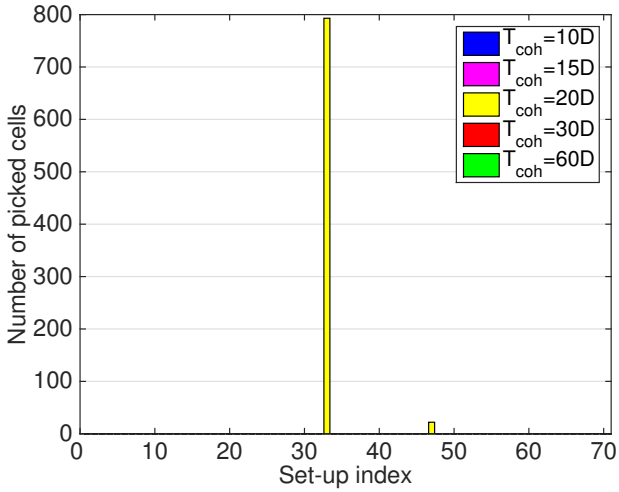
(a) G347.3 (close and young for Vela Jr.)



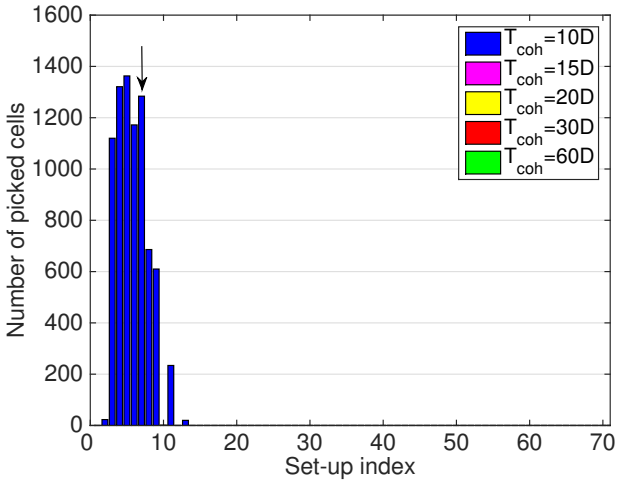
(b) G347.3 (close and old for Vela Jr.)



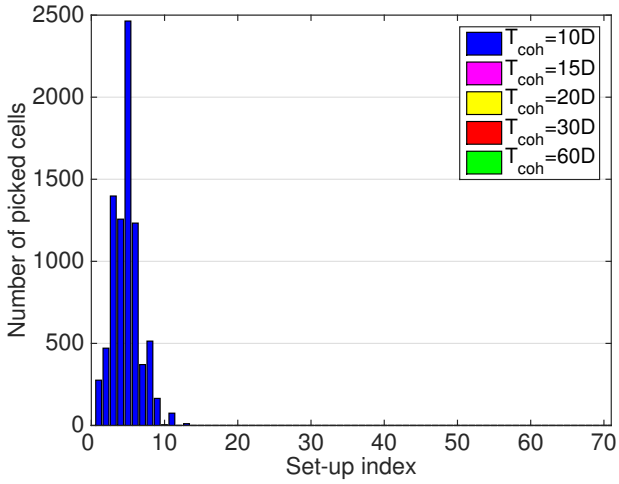
(c) Vela Jr. (close and young for Vela Jr.)



(d) Vela Jr. (close and old for Vela Jr.)



(e) Cas A (close and young for Vela Jr.)



(f) Cas A (close and old for Vela Jr.)

FIG. 5: Optimal set-up details. The left hand-side figures show the set-ups assuming Vela Jr. is close and young. The right hand-side figures show the set-ups assuming Vela Jr. is close and old. As explained in in Section V D, the arrows indicate the set-ups that will finally be adopted in the search.

TABLE IV: Effect of using the wrong age prior for Vela Jr.

	uniform distribution on f, $\dot{f}$		log-uniform distribution on f, $\dot{f}$		
	C for Vela Jr.	F for Vela Jr.	C for Vela Jr.	F for Vela Jr.	$\langle loss \rangle$
$\mathcal{R}$ if Vela Jr. is 4300 yrs old	14.8%(16.1%):	7.9%(8.6%):	6.4%(11.4%):	4.7%(4.9%):	
but prior assumes 700 yrs of age	$loss$ =**7.9%**	$loss$ =**9.0%**	$loss$ =**43.9%**	$loss$ =**3.7%**	**16.1%**
$\mathcal{R}$ if Vela Jr. is 700 yrs old	6.1%(18.7%):	5.1%(11.4%):	4.3%(9.2%):	3.0%(4.5%):	
but prior assumes 4300 yrs of age	$loss$ =**67.4%**	$loss$ =**55.1%**	$loss$ =**53.5%**	$loss$ =**33.3%**	**52.3%**

Notes: The first number in each table cell is the $\mathcal{R}$ assuming a mismatch between the prior used in the optimization and the true age of the object. The number in parenthesis is the $\mathcal{R}$ if the prior is matched to the age of the object. The first number is always smaller than the $\mathcal{R}$ in parenthesis and the ratio in detection probability is in bold font.

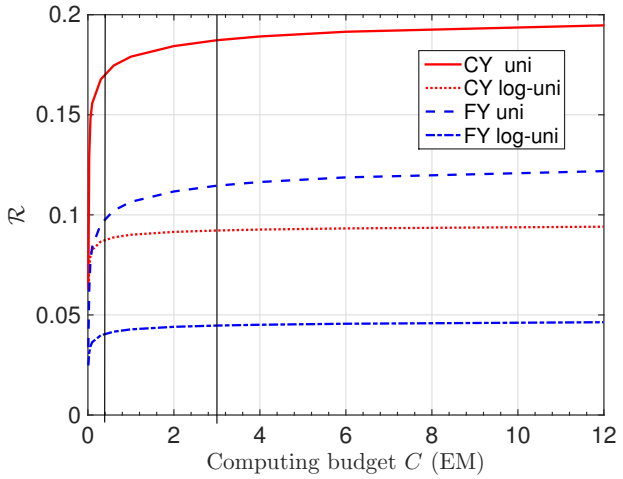


FIG. 6: $\mathcal{R}$ versus computing budget C

D. Simplifications

As shown in the previous Sections, the optimal search set-up may well comprise different coherent time baselines $T_{\rm coh}$ in different $f-\dot{f}$ ranges for each target, and different grid spacings for the same $T_{\rm coh}$. For instance, under the CY and uniform priors assumption for Vela Jr., the set-prescription comprises 2 different set-ups for $T_{\rm coh}$ = 10 days, 7 set-ups for $T_{\rm coh}$ = 15 days, and 5 set-ups for $T_{\rm coh}$ = 20 days for Vela Jr.; 11 set-ups for $T_{\rm coh}$ = 10 days for Cas A and 27 set-ups distributed over four different $T_{\rm coh}$ values (10, 15, 20, 30 days) for G347.3.

The analysis of the results from a search comprising this diversity in set-ups and coherent time baselines is quite daunting. So we examine the following question: how much detection probability would be lost if i) we considered only a single set-up per target ii) we extended the search frequency range for all targets to be fixed between 20-1500 Hz?

Limiting the set-up for each target to the 71 set-ups considered in Section V A we have $71^3 = 357911$ combinations of setups. The computing costs of these ranges from a few 0.1EM to a few hundreds EM. We want a computing budget of a few EM. In Fig. 7 we zoom in the cost range from 0 to 6EM and determine the total detection probability for all the set-up combinations.

TABLE V: Profile of the chosen set-ups

Targets	m_f	$m_{\dot{f}}$	$m_{\ddot{f}}$	$\gamma^{(1)}$	$\gamma^{(2)}$	$T_{\rm coh}$	$\langle\mu\rangle$
Cas A	0.3	0.5	0.003	4	20	10D	41.2%
Vela Jr.	0.15	0.3	0.003	8	20	15D	15.8%
G347.3	0.15	0.2	0.003	8	10	20D	12.1%

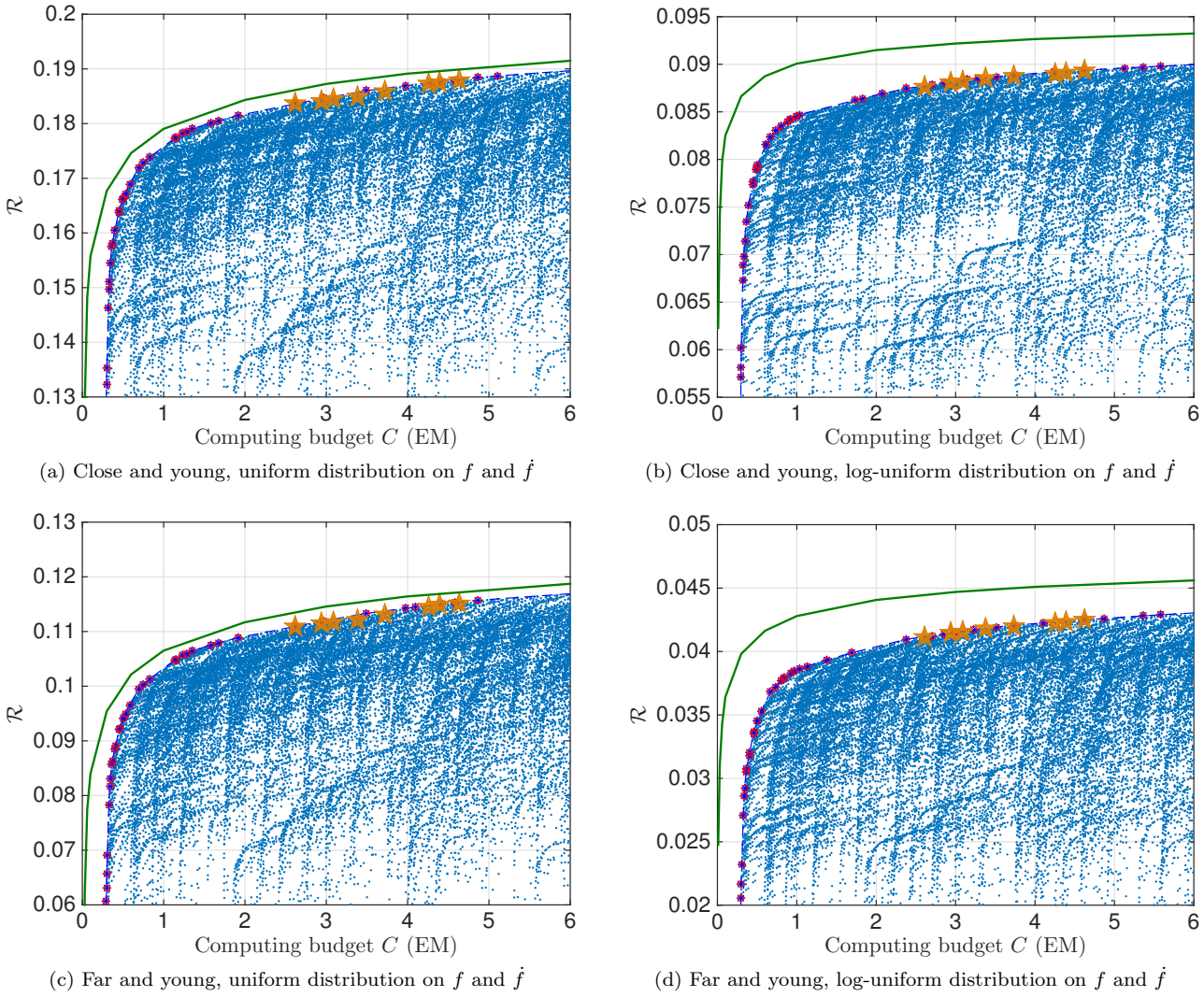
We find that the loss due to having restricted the choice to a single set-up per target is smaller than $\sim$0.3% across all priors. Since this loss is much smaller than the total detection probability, we are persuaded to adopt this simplification and save ourselves a great deal of effort in the results post-processing phase.

We pick a computing budget just below 5 EM and hence the set-ups corresponding to the right-most golden star of Fig. 7. Actually, all these golden stars correspond to the same set-ups independently of the priors for both the distance of Vela Jr. and f and $\dot{f}$. The details of the chosen set-up are listed in Table V.

We use arrows to indicate the three chosen set-ups in Fig. 5. Under the CY and uniform $f, \dot{f}$ assumption for Vela Jr., the chosen set-up for G347.3 is also the dominant one from the optimization scheme; the chosen set-up for Cas A is the second dominant one from the optimization scheme. Although the chosen set-up for Vela Jr. is the fourth dominant one from the optimization procedure (see Fig. 5(c)) the chosen $T_{\rm coh}$ = 15 days is the same that of the dominant set-up from the optimization scheme. These differences can be explained considering that we take different prior combinations into account overall and the computing budget set for Fig. 5 is 3 EM which is below the final-set budget 5 EM.

VI. CONCLUSIONS

Following the search optimization procedure proposed in [22], we design a search using a few months of Einstein@Home optimized for a data set like the LIGO O1 data. We concentrate on 3 targets: Vela Jr., Cas A



(a) Close and young, uniform distribution on f and $\dot{f}$

(b) Close and young, log-uniform distribution on f and $\dot{f}$

(c) Far and young, uniform distribution on f and $\dot{f}$

(d) Far and young, log-uniform distribution on f and $\dot{f}$

FIG. 7: $\mathcal{R}$ versus computing budget C from 357911 combinations of setups with different priors. Each dot on this plot is a combination of setups, one per target. The green curve shows the optimal $\mathcal{R}$ as a function of C. The golden stars highlight the best setup combinations around 3 EM.

and G347.3. We extend the method proposed in [22] by adding more dimensions to the optimization : i.e. we consider different search set-ups for the same coherent time baseline $T_{\rm coh}$, varying the template banks in frequency and spindown, and we fold-in the *measured* mismatch distributions from the different banks.

We also investigate how a mistake in choice of the astrophysical prior on the age of the target that contributes the most to the detection probability (Vela Jr.) would impact the detection probability, and then pick the prior that minimizes the loss.

We study the dependency of the attainable detection probability on the computing budget, and, within practical constraints from running the search, we make sure that we have nearly saturated the detection probability growth. We pick a computing budget of ~ 5 EM.

After having obtained the optimal combination of setups for the different targets in the different regions of parameter space, we significantly simplify it in order to make the post-processing of the results less cumbersome. Even limiting the search set-up to a single one per target, we are able to achieve this without significant degradation in the detection probability. One may wonder if this doesn't prove that the optimization scheme is actually not very important. In a sense it does, at least for this data set. However, without knowing what the optimal is, we would not have been able to judge the goodness of any empirically motivated set-up.

This is the final set-up chosen:

- For the youngest source Cas A, a set-up with 10 days coherent time baseline (12 segments) will be used. The computing cost employed on searching for a signal from Cas A is 1.7 EM. The detection probability is 1.2%, if we assume uniform priors in

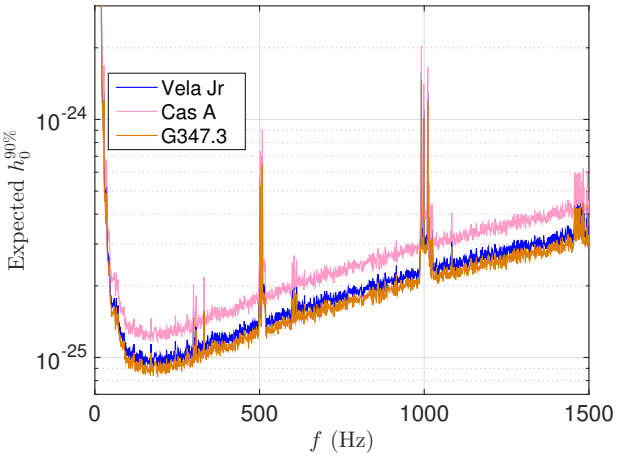


FIG. 8: Expected 90% confidence upper limit on the GW amplitude $h_0^{90\%}$ from the directed searches proposed here, based on the noise level of the LIGO O1 data and on the estimated sensitivity of the proposed searches. We stress that this is *not* an observational result but a prediction of it.

frequency and spin-down; the detection probability is 0.2%, if we assume log-uniform priors.

- For the closest source Vela Jr., a set-up with 15 days coherent time baseline (8 segments) will be used. The computing cost employed on searching for a signal from Vela Jr. is 2.2 EM. If we assume uniform priors in frequency and spin-down, a distance of 200 pc and an age of 700 yrs, the detection probability is 14.5%. The detection probability drops to 3.8% if we assume a distance of 750 pc and an age of 4300 yrs. If we assume log-uniform priors in frequency and spin-down, a distance of 200 pc and an age of 700 yrs, the detection probability is 7.4%. The detection probability drops to 3.0% if we assume a distance of 750 pc and an age of 4300 yrs.

- For the second closest source G347.3, a set-up with 20 days coherent time baseline (6 segments) will be used. The computing cost employed on searching for a signal from Vela Jr. is 0.7 EM. The detection probability is 3.1%, if we assume uniform priors in frequency and spin-down; the detection probability is 1.4%, if we assume log-uniform priors.

The search that we propose here is a deep and broad frequency search for all the three targets Vela Jr., Cas A and G347.3. It was set-up and ran for a few months on the volunteer computing project Einstein@Home during the first half of 2017. The post-processing of the results will be reported in a separate paper. In case no signal is detected the expected 90% confidence upper limit on the GW strain amplitude $h_0^{90\%}$ is shown in Fig. 8. At the detector's most sensitive frequencies, ≈150 Hz, $h_0^{90\%}$ for Cas A is 1.4×10^{-25}, for Vela Jr. 1.0×10^{-25} and for G347.3 9×10^{-26}. We note that $h_0^{90\%} = 1.4 \times 10^{-25}$ for Cas A is 2 times smaller than the S6 upper limit [9] and deep searches on Vela Jr. and G347.3 have never been done yet. This can be quantified by using the notion of sensitivity depth defined in [50]. The sensitivity depth $\mathcal{D}^{90\%}$ of this search is 61.5 $\mathrm{Hz}^{-1/2}$ for Cas A, 79.1 $\mathrm{Hz}^{-1/2}$ for Vela Jr., and 85.8 $\mathrm{Hz}^{-1/2}$ for G347.3.

VII. ACKNOWLEDGMENTS

J.M. acknowledges support by the IMPRS on Gravitational Wave Astronomy at the Max Planck Institute for Gravitational Physics in Hannover. M.A.P. gratefully acknowledges support from NSF PHY grant 1104902. We thank David Keitel, Ra Inta and Ben Owen for valuable comments. This paper was assigned LIGO document number P1700166.

[1] M. A. Papa *et al.*, Phys. Rev. **D94**, 122006 (2016), arXiv:1608.08928 [astro-ph.IM].
[2] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al., Phys. Rev. D **94**, 042002 (2016), arXiv:1605.03233 [gr-qc].
[3] B. P. Abbott *et al.* (Virgo, LIGO Scientific), (2017), arXiv:1701.07709 [astro-ph.HE].
[4] B. Abbott *et al.* (LIGO Scientific Collaboration), Astrophys.J. **683**, L45 (2008), arXiv:0805.4758 [astro-ph].
[5] J. Abadie *et al.* (LIGO Scientific Collaboration, Virgo Collaboration), Astrophys.J. **737**, 93 (2011), arXiv:1104.2712 [astro-ph.HE].
[6] J. Aasi *et al.* (VIRGO, LIGO Scientific), Phys. Rev. **D91**, 022004 (2015), arXiv:1410.8310 [astro-ph.IM].
[7] J. Aasi *et al.* (VIRGO, LIGO Scientific), Phys. Rev. **D93**, 042007 (2016), arXiv:1510.03621 [astro-ph.IM].
[8] LIGO Scientific Collaboration and Virgo Collaboration, `https://dcc.ligo.org/P1700127` (2017).
[9] S. J. Zhu, M. Papa, H.-B. Eggenstein, R. Prix, K. Wette, B. Allen, O. Bock, D. Keitel, B. Krishnan, B. Machenschalk, M. Shaltev, and X. Siemens, ArXiv e-prints (2016), arXiv:1608.07589 [gr-qc].
[10] J. Aasi *et al.* (LIGO), Astrophys. J. **813**, 39 (2015), arXiv:1412.5942 [astro-ph.HE].
[11] J. Aasi *et al.* (VIRGO, LIGO), Phys. Rev. **D93**, 042006 (2016), arXiv:1510.03474 [gr-qc].
[12] T. D. Abbott *et al.* (VIRGO, LIGO), (2016), arXiv:1607.02216 [gr-qc].
[13] G. D. Meadors, E. Goetz, K. Riles, T. Creighton, and F. Robinet, (2016), arXiv:1610.09391 [gr-qc].
[14] J. Aasi *et al.* (VIRGO, LIGO Scientific), Phys. Rev. **D88**, 102002 (2013), arXiv:1309.6221 [gr-qc].

[15] P. R. Brady and T. Creighton, Phys. Rev. D **61**, 082001 (2000), arXiv:gr-qc/9812014.

[16] B. Krishnan, A. M. Sintes, M. A. Papa, B. F. Schutz, S. Frasca, *et al.*, Phys. Rev. D **70**, 082001 (2004), arXiv:gr-qc/0407001.

[17] M. A. Papa, B. F. Schutz, and A. M. Sintes, in *Gravitational waves: A challenge to theoretical astrophysics. Proceedings, Trieste, Italy, June 6-9, 2000* (2000) pp. 431–442, arXiv:gr-qc/0011034 [gr-qc].

[18] B. Krishnan (LIGO Scientific Collaboration), Class. Quant. Grav. **22**, S1265 (2005), arXiv:gr-qc/0506109.

[19] C. Cutler, I. Gholami, and B. Krishnan, Phys. Rev. D **72**, 042004 (2005), arXiv:gr-qc/0505082.

[20] H. J. Pletsch and B. Allen, Phys. Rev. Lett. **103**, 181102 (2009), arXiv:0906.0023 [gr-qc].

[21] H. J. Pletsch, Phys. Rev. D **78**, 102005 (2008), arXiv:0807.1324 [gr-qc].

[22] J. Ming, B. Krishnan, M. A. Papa, C. Aulbert, and H. Fehrmann, Phys. Rev. D **93**, 064011 (2016), arXiv:1510.03417 [gr-qc].

[23] R. Prix and M. Shaltev, Phys.Rev. **D85**, 084010 (2012), arXiv:1201.4321 [gr-qc].

[24] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, and et al., Physical Review Letters **116**, 131103 (2016), arXiv:1602.03838 [gr-qc].

[25] G. G. Pavlov, D. Sanwal, B. Kızıltan, and G. P. Garmire, The Astophysical Journal Letters **559**, L131 (2001), astro-ph/0108150.

[26] J. Aasi, B. P. Abbott, R. Abbott, T. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, T. Adams, and et al., ArXiv e-prints (2014), arXiv:1412.5942 [astro-ph.HE].

[27] H. Tananbaum, IAU Circ. **7246** (1999).

[28] J. E. Reed, J. J. Hester, A. C. Fabian, and P. F. Winkler, Astrophys. J. **440**, 706 (1995).

[29] R. A. Fesen, M. C. Hammell, J. Morse, R. A. Chevalier, K. J. Borkowski, M. A. Dopita, C. L. Gerardy, S. S. Lawrence, J. C. Raymond, and S. van den Bergh, Astrophys. J. **645**, 283 (2006), astro-ph/0603371.

[30] G. Cassam-Chenaï, A. Decourchelle, J. Ballet, J.-L. Sauvageot, G. Dubner, and E. Giacani, Astronomy and Astrophysics **427**, 199 (2004), astro-ph/0407333.

[31] Z. R. Wang, Q.-Y. Qu, and Y. Chen, Astronomy and Astrophysics **318**, L59 (1997).

[32] R. P. Mignani, S. Zaggia, A. de Luca, R. Perna, N. Bassan, and P. A. Caraveo, Astronomy and Astrophysics **484**, 457 (2008), arXiv:0803.3722.

[33] Details of the Einstein@Home project can be found at `http://einsteinathome.org`.

[34] P. Jaranowski, A. Królak, and B. F. Schutz, Phys. Rev. D **58**, 063001 (1998), arXiv:gr-qc/9804014.

[35] C. Cutler and B. F. Schutz, Phys. Rev. D **72**, 063006 (2005), arXiv:gr-qc/0504011.

[36] H. J. Pletsch, Phys. Rev. D **82**, 042002 (2010), arXiv:1005.0395 [gr-qc].

[37] LAL/LALapps Software Suite.`http://www.lsc-group.phys.uwm.edu/daswg/projects/lalsuite.html`.

[38] P. Patel, X. Siemens, R. Dupuis, and J. Betzwieser, Phys. Rev. D **81**, 084032 (2010), arXiv:0912.4255 [gr-qc].

[39] R. Prix, *Resampling FFT-based implementation of the F-statistic*, Tech. Rep. **T1600531** (LIGO, 2017) `https://dcc.ligo.org/LIGO-T1600531-V3`.

[40] D. Keitel, R. Prix, M. A. Papa, P. Leaci, and M. Siddiqi, Phys. Rev. D **89**, 064023 (2014), arXiv:1311.5738 [gr-qc].

[41] D. Keitel and R. Prix, Classical and Quantum Gravity **32**, 035004 (2015), arXiv:1409.2696 [gr-qc].

[42] D. Keitel, Phys. Rev. D **93**, 084024 (2016), arXiv:1509.02398 [gr-qc].

[43] `http://www.aei.mpg.de/24838/02_Computing_and_ATLAS`.

[44] G. E. Allen, K. Chow, T. DeLaney, M. D. Filipovic, J. C. Houck, T. G. Pannuti, and M. D. Stage, Astrophys. J. **798**, 82 (2015), arXiv:1410.7435 [astro-ph.HE].

[45] J. Abadie *et al.* (LIGO Scientific Collaboration), Astrophys. J. **722**, 1504 (2010), arXiv:1006.2535 [gr-qc].

[46] C. Horowitz and K. Kadau, Phys. Rev. Lett. **102**, 191102 (2009), arXiv:0904.1986 [astro-ph.SR].

[47] N. Andersson, V. Ferrari, D. Jones, K. Kokkotas, B. Krishnan, *et al.*, Gen.Rel.Grav. **43**, 409 (2011), arXiv:0912.0384 [astro-ph.SR].

[48] S. Shapiro and S. Teukolsky, *Black holes, white dwarfs, and neutron stars* (Wiley-VCH Verlag GmbH & Co. KgaA, Weinheim, 2004).

[49] J. Traub and H. Woniakowski, Operations Research Letters **1**, 59 (1982).

[50] B. Behnke, M. A. Papa, and R. Prix, Phys. Rev. **D91**, 064007 (2015), arXiv:1410.5997 [gr-qc].