The Hough transform search for continuous gravitational waves

Badri Krishnan,[1,*] Alicia M. Sintes,[1,2,†] Maria Alessandra Papa,[1,‡]
Bernard F. Schutz,[1,3,§] Sergio Frasca,[4,5,¶] and Cristiano Palomba[5,**]

[1] *Max-Planck-Institut für Gravitationsphysik, Albert-Einstein-Institut, Am Mühlenberg 1, D-14476 Golm, Germany*
[2] *Departament de Física, Universitat de Les Illes Balears, Cra. Valldemossa Km. 7.5, E-07122 Palma de Mallorca, Spain*
[3] *Cardiff University, Cardiff, CF2 3YB, United Kingdom*
[4] *Dip. di Fisica, Università di Roma "La Sapienza" P.le A. Moro 2, I-00185 Rome, Italy*
[5] *INFN Sezione di Roma, P.le A. Moro 2, I-00185 Rome, Italy*

(Dated: Draft of May 31, 2004)

This paper describes an incoherent method to search for continuous gravitational waves. The method is based on the Hough transform, a well known technique used for detecting patterns in digital images. We apply the Hough transform to detect patterns in the time-frequency plane of the data produced by an earth-based gravitational wave detector. Two different flavours of searches will be considered, depending on the type of input to the Hough transform: either Fourier transforms of the detector data or the output of a coherent matched-filtering type search. We present the technical details for implementing the Hough transform algorithm for both kinds of searches, the statistics of the methods and their sensitivity.

I. INTRODUCTION

Rapidly rotating neutron stars are expected to be the primary sources of continuous gravitational waves, and the current generation of earth-based gravitational wave detectors might be able to detect them. Recent analysis of data from the first science runs of the LIGO [1, 2] and GEO [3, 4] interferometric detectors has already led to upper limits on the gravitational wave emitted by the pulsar J1939+2134 and its equatorial ellipticity [5]. The analysis of future science runs is expected to lead to upper limits below other astrophysical constraints, and eventually to detections.

The analyses presented in [5] were based on the coherent integration of the detectors' output for the entire observation time (approximately 17 days) and used a Bayesian time-domain method and a frequentist frequency-domain [6] approach. The searches were not computationally expensive, targeting a single known pulsar and processing only a narrow frequency band of about 0.5 Hz around the pulsar emission frequency for a fixed sky location and spindown rate known from radio observations.

Future continuous wave searches will involve searching longer data stretches (of order weeks to months) for unknown sources over a large frequency band, vast portions of the sky and spindown parameter values. It is well known that the computational cost of coherent techniques for searches of this type is absolutely prohibitive [7]. Thus hierarchical methods have been proposed. In hierarchical strategies incoherent techniques (less sensitive and less computationally expensive) are used to scan the data and the parameter space for interesting candidates which are then followed up with coherent searches. Different strategies can be envisaged that combine the data incoherently. All methods involve using in some way the power from the Fourier transforms of short stretches of data: in the frequency bins where the signal is present there should systematically be an excess of power. In order to compensate for the frequency modulation imposed on the signal by the Earth's motion and the pulsar's spindown during the observation period, one must use not the power from the same frequency bins in each successive Fourier transform, but rather from the bins where one expects the signal peak to be. In the so called stack-slide method, one "slides" the frequency bins of each Fourier transform to line-up the signal peaks and then simply sums the power [8]. The Hough transform method can be seen as a variation on this where, after the sliding, one sums not the power but just zeros and ones, depending on whether the power in the frequency bin meets some criterion. Whereas in low signal-to-noise conditions in Gaussian noise the standard power summing method is possibly optimal we expect the Hough transform method to be more robust in the presence of large spectral disturbances.

The Hough transform is a robust parameter estimator of multi-dimensional patterns in images and it finds many applications in astronomical data analysis. It was initially developed by Paul Hough to analyze bubble chamber pictures at CERN, and later patented by IBM [10, 11]. It is currently being used to analyze data from the LIGO and GEO detectors. The codes employed for these analyses are freely available as part of the LIGO Algorithms Library [12]. The VIRGO project [13, 14] is also setting up a similar hierarchical search pipeline. Studies of hierarchical strategies can be found in [15, 16, 17, 18, 19].

This paper is organized as follows: section II briefly

*Electronic address: badri.krishnan@aei.mpg.de
†Electronic address: alicia.sintes@uib.es
‡Electronic address: maria.alessandra.papa@aei.mpg.de
§Electronic address: schutz@aei.mpg.de
¶Electronic address: sergio.frasca@roma1.infn.it
**Electronic address: cristiano.palomba@roma1.infn.it

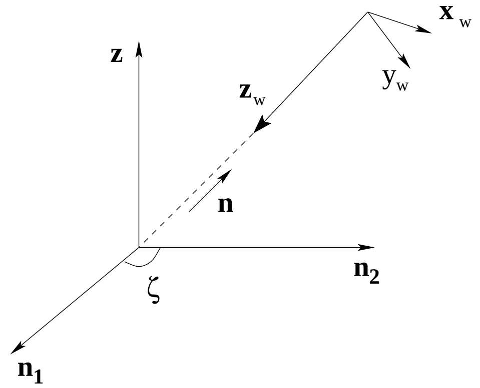

FIG. 1: The detector frame and the wave frame.

describes the expected waveforms from an isolated spinning neutron star and summarizes the general strategy of a hierarchical search. Section III presents the general idea of the Hough transform and section IV describes its implementation for non-demodulated input data, and section V studies its statistical properties. Section VI describes the Hough search using demodulated input data and finally section VII summarizes our main results.

II. PRELIMINARIES

A. The signal from a pulsar

In this section we fix our notation and briefly review the expected gravitational wave signal from a spinning neutron star. Further details about the pulsar signal can be found in [6]; a concise review of the possible physical mechanisms that may be causing pulsars to emit gravitational waves can be found in [5]. For our purposes, we only need the form of the gravitational-wave signal as seen by an Earth based detector.

Let $\mathbf{n}_1$ and $\mathbf{n}_2$ denote the unit vectors pointing along the arms of the detector and denote by ζ the angle between the arms. Let $\mathbf{z}$ be the unit vector parallel to $\mathbf{n}_1 \times \mathbf{n}_2$. Apart from the detector frame $(\mathbf{n}_1, \mathbf{n}_2, \mathbf{z})$, we also have the wave frame $(\mathbf{x}_w, \mathbf{y}_w, \mathbf{z}_w)$ in which the unit vector $\mathbf{z}_w$ is along the direction of propagation of the wave and $(\mathbf{x}_w, \mathbf{y}_w, \mathbf{z}_w)$ form a right-handed orthonormal system. Finally, $\mathbf{n} = -\mathbf{z}_w$ is the unit vector pointing in the direction of the neutron star; see figure 1. The spacetime metric $g_{\mu\nu}$ can be written as a perturbation of the flat metric $\eta_{\mu\nu}$: $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$. The received gravitational wave $h_{\mu\nu}$ has the form

$$h_{\mu\nu}(t) = h_+(t)\,(\mathbf{e}_+)_{\mu\nu} + h_\times(t)\,(\mathbf{e}_\times)_{\mu\nu} \tag{2.1}$$

where $\mathbf{e}_+ = \mathbf{x}_w \otimes \mathbf{x}_w - \mathbf{y}_w \otimes \mathbf{y}_w$ and $\mathbf{e}_\times = \mathbf{x}_w \otimes \mathbf{y}_w + \mathbf{y}_w \otimes \mathbf{x}_w$, and t denotes clock time at the location of the (moving, accelerating) detector, which we refer to as *detector time*. The waveforms for the two polarizations are

$$h_+(t) = A_+ \cos\Phi(t)\,, \qquad h_\times(t) = A_\times \sin\Phi(t) \tag{2.2}$$

where $\Phi(t)$ is the phase of the gravitational wave and $A_{+,\times}$ are the amplitudes; $A_{+,\times}$ are constant in time and depend on the other pulsar parameters such as its rotational frequency, moments of inertia, the orientation of its rotation axis, its distance from Earth etc. The phase $\Phi(t)$ takes its simplest form when the time coordinate used is $t_{\rm ns}$, the proper time in the rest frame of the neutron star:

$$\Phi_{\rm ns}(t_{\rm ns}) = \phi_0 + 2\pi \sum_{n=0}^{\infty} \frac{f_{(n)}^{(\rm ns)}}{(n+1)!} t_{\rm ns}^{n+1} \tag{2.3}$$

where ϕ_0, $f_{(0)}^{(\rm ns)}$ and $f_{(n)}^{(\rm ns)}$ are respectively the phase, instantaneous frequency and the spindown parameters in the rest frame of the star at the fiducial start time $t_{\rm ns} = 0$.

We refer the reader to [6] for the expression of $\Phi(t)$ in the detector frame as a function of detector time. For our purposes, we only need to know that the instantaneous frequency $f(t)$ of the wave as observed by the detector is given, to a very good approximation, by the familiar non-relativistic Doppler formula:

$$f(t) - \hat{f}(t) = \hat{f}(t)\frac{\mathbf{v}(t)\cdot\mathbf{n}}{c} \tag{2.4}$$

where $\mathbf{v}(t)$ is the detector velocity in the solar system frame and $\hat{f}$ is given by

$$\hat{f}(t) = f_{(0)} + \sum_{n=1}^{\infty} \frac{f_{(n)}}{n!}(t - t_0)^n\,. \tag{2.5}$$

The time t_0 is the fiducial detector time at the start of the observation and the $f_{(n)}$ are the spindown parameters as measured in the solar system frame (these need not be equal to the $f_{(n)}^{(\rm ns)}$; see [6]). We have assumed that the neutron star is moving with uniform speed relative to the Sun and is so far away that there are no observable proper-motion effects. (These could be taken into account if necessary, at the cost of introducing further parameters.)

The detector output will be a linear combination of h_+ and $h_\times$:

$$h(t) = F_+(\mathbf{n},\psi)\,h_+(t) + F_\times(\mathbf{n},\psi)\,h_\times(t) \tag{2.6}$$

where $F_{+,\times}$ are known as the the antenna pattern functions of the detector and depend on the direction $\mathbf{n}$ to the star and also on the polarization angle ψ which determines the orientation of the $(\mathbf{x}_w, \mathbf{y}_w)$ axes in their plane. In addition, the antenna pattern functions also depend on the detector parameters such as the angle ζ between the arms, the length of the arms etc. Due to the motion of the Earth, $F_{+,\times}(\mathbf{n},\psi)$ depend implicitly on time and for notational convenience, we shall usually denote the antenna pattern functions as $F_{+,\times}(t)$. Thus, the received signal is both amplitude- and frequency-modulated.

The search method described in this paper depends on finding a signal whose frequency evolution fits the pattern produced by the Doppler shift and the spindown. The search takes place in the time-frequency plane formed by successive spectra, in each of which the frequency and amplitude of the received signal are effectively constant. The parameters which determine this pattern are the ones which appear in equation (2.4), namely, $(f_{(0)}, \{f_{(n)}\}, \mathbf{n})$; these parameters will be collectively denoted by the symbol $\vec{\xi}$.

The amplitudes $A_{+,\times}$ are determined by the other pulsar parameters such orientation of its axis, its ellipticity, its distance from Earth etc. The search method presented in this paper depends only on the phase model of equation (2.3). The exact form of the amplitudes is model dependent. As an illustrative example, if we consider the wave emitted by a deformed spinning neutron star as in [5], if f_r is the rotational frequency of the star, the frequency of the gravitational wave is $2f_r$. The additional parameters determining this component of the pulsar signal are ι and h_0 where where ι is the angle between the pulsar's axis of rotation and the vector $\mathbf{n}$, and h_0 characterizes the amplitude of the emitted gravitational wave. The amplitudes $A_{+,\times}$ are:

$$A_+ = \frac{1}{2}h_0(1+\cos^2\iota)\,, \qquad (2.7)$$

$$A_\times = h_0\cos\iota\,. \qquad (2.8)$$

If we assume the emission mechanism is due to deviations of the pulsar's shape from perfect axial symmetry, then the amplitude h_0 is given by

$$h_0 = \frac{16\pi^2 G}{c^4}\frac{I_{zz}\epsilon f_r^2}{d} \qquad (2.9)$$

where d is the distance of the star from Earth, I_{zz} is the z-z component of the star's moment of inertia with the z-axis being its axis of symmetry, and $\epsilon := (I_{xx}-I_{yy})/I_{zz}$ is the equatorial ellipticity of the star.

B. A multi-stage hierarchical search

Consider performing a blind search for pulsars using a bank of templates and relying only on coherent matched filter techniques. Since a larger observation time implies better resolution in the space of frequencies, spindowns and sky-positions, the number of templates increases rapidly as a function of the total observation time. Consider for example, an all-sky search for young-fast pulsars, i.e. for hypothetical signals with frequency $\hat{f} < f_{\rm max} = 1000$Hz and spindown ages greater than $\tau > \tau_{\rm min} = 40$yr. ($\tau := \hat{f}/f_{(1)}$.) Let s be the number of spindown parameters that we search over and let $T_{\rm obs}$ be the total observation time. The number of templates required for this search has been calculated in [7]:

$$N_p \approx \max_{s\in\{0,1\ldots\}}[N_s F_s(T_{\rm obs})] \qquad (2.10)$$

where

$$N_s = \left(\frac{f_{\rm max}}{1\rm kHz}\right)^{s+2}\left(\frac{40\rm yr}{\tau_{\rm min}}\right)^{s(s+1)/2}\,. \qquad (2.11)$$

gives the spindown scaling and F_s is a function that depends on the observation time; for large observation time, $F_s \propto T_{\rm obs}^5$. If we take the maximum allowed fractional mismatch in observed signal power between the signal and the template to be 0.3, then for, say, $s=2$:

$$F_2(T_{\rm obs}) \approx 2\times 10^7 \times \left(\frac{T_{\rm obs}}{1\rm day}\right)^5\,. \qquad (2.12)$$

Thus, even for a 10 day search over two spindown parameters, $N_p \approx 2\times 10^{12}$. The computational requirements for a search over these many templates is also estimated in [7]. It turns out that for the 10 day long search, if we wish to analyze the data in roughly real time, we would require a computational power of $\sim 10^{17}$ Flops. This is roughly 5 or 6 orders of magnitude greater than presently available computational resources. Even if we insisted on searching over only a single spindown parameter, the computational requirement would turn out to be $\sim 10^{14}$ Flops which is still too large even for an observation time of just 10 days. *We therefore conclude that a search over any significant portion of parameter space for unknown pulsars is not possible in the foreseeable future if we restrict ourselves to fully coherent methods.*

One possible way to perform such a blind search would be to exploit the fact that N_p increases faster than linearly with $T_{\rm obs}$. Thus if we break up the data set into smaller segments, it might be feasible to analyze each data segment coherently. An incoherent method is then used as a computationally inexpensive and sub-optimal way of combining the outputs of the different coherent segments. This would be one step in a multi-stage hierarchical scheme; see figure 2.

In this scheme, one starts with a data stream covering a total observation time $T_{\rm obs}$. The available data is divided into smaller segments and each segment analyzed coherently, i.e. the phase information within a segment is fully taken into account. The results of the coherent analysis of the different segments are then combined incoherently. The output of the incoherent step is a set of possible pulsar candidates. If necessary, acquire fresh data and repeat the above procedure analyzing only the candidates selected by the previous step. Once this procedure has been iterated the desired number of times and the number of candidates in parameter space is small enough, the candidates are analyzed by using the entire data stream coherently. The final output of the search is, of course, either a detection or an upper limit.

The exact number of times the incoherent step must be repeated and the thresholds that one must set at each stage are decided by optimizing the sensitivity subject to the obvious constraints on the desired signal strength we wish to detect, the desired confidence level and the

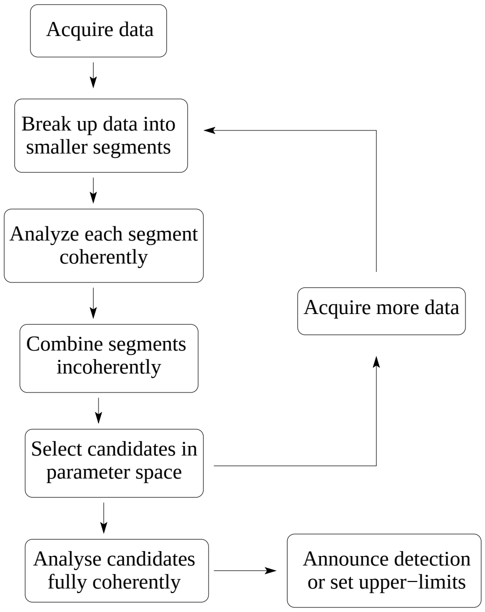

FIG. 2: A hierarchical scheme for analysis large parameter space volumes for CW searches. Each step only analyses the regions in parameter space that have not been discarded by any of the previous steps.

amount of total data available. Preliminary investigations of this optimization are reported in [8] and more detailed results will be presented elsewhere [20].

Hierarchical searches like this are typically effective only when looking for signals that, in the final coherent search over the whole data set, have relatively high signal-to-noise ratio. The method only works if the incoherent step succeeds in reducing the number of points in parameter space that one must search over. A signal that is only, say, at two-sigma in the final step will be too weak in the initial shorter coherent transforms to be selected by any criterion that would eliminate other ("pure noise") parameter points. Signals must be strong enough so that they consistently bias upwards the power in the coherent stages. Remarkably, this does not actually reduce the sensitivity of a hierarchical search over the hypothetical fully coherent search we described above. The reason is that the size of the parameter space is so large that, even in a fully coherent search, signals must be unusually strong in order to be detected with enough significance to be recognized. In our case, a fully coherent signal-to-noise ratio of 10 or more is needed for a significant detection over a period of several months, and we will see in equation (5.40) below and the subsequent discussion that our incoherent methods do worse than this by factors of between two and 5, while permitting much larger regions of parameter space to be surveyed.

Finally we mention one important detail to be used later, namely, the nature of the coherent analysis of each data segment and of the final follow-up stage. In this paper we consider two possible alternatives. The first alternative is just to use the Fourier transform of data segments that are so short that no frequency modulation or spindown is measurable. These transforms are called SFTs (Short time base-line Fourier Transform) and may represent up to 30 minutes of data. The candidates for the incoherent step are selected based on the *normalized* SFT power, i.e. on the power divided by the noise floor estimate.

If longer coherent stages are required for better sensitivity, then one must use *demodulated* data, i.e. remove the effects of Earth's spin and orbital motion and also of the pulsar spindown. This demodulation must be done separately for different regions of the sky and spindown parameter space, but it also brings in other parameters, such as the polarization angle ψ, because of the effects of amplitude modulation. These extra parameters, which are not part of our Hough-transform search space, can be eliminated by requiring the coherent stage to produce the $\mathcal{F}$ statistic described in [6] and used in [5] for analyzing the data from the first science runs of the LIGO and GEO detectors. In this case, we would select frequency bins based on the value of the $\mathcal{F}$ statistic. The search based on SFTs will be called the *non-demodulated* search and is described in sections IV and V. The search using the $\mathcal{F}$ statistic is the search with *demodulated* data and is described in section VI.

III. THE HOUGH TRANSFORM

As mentioned in the introduction, the Hough transform is a robust parameter estimator for patterns in digital images. It can be used to identify the parameter(s) of a curve which best fits a set of given points. The idea is to map data into the parameter space, which is appropriately quantized, and then seek for the most likely parameter values. The Hough Transform provides robustness against discontinuous or missing data. The Hough Transform has become in the last two decades a standard tool in the domain of artificial vision for the recognition of patterns that can be parameterized like straight lines, polynomials, circles, etc., in a suitable parameter space.

For our purposes, a pattern is a collection $\mathcal{C}$ of smooth hypersurfaces [24] in some differentiable manifold M. Assume that there is a manifold Σ of parameters which describes elements of $\mathcal{C}$; i.e. there exists a function $f : \Sigma \to \mathcal{C}$ providing a one-one association between points on Σ and elements of $\mathcal{C}$ (see figure.

A simple example is the case when M is $\mathbb{R}^2$ with coordinates x and y, and $\mathcal{C}$ is the collection of straight lines in this (x, y) plane. Since all straight lines are described by an equation of the form $y = mx + c$, the parameter space Σ is also $\mathbb{R}^2$, with coordinates (m, c) – the slope and the y-intercept of the straight lines. The function f maps the point (m, c) to the straight line $y = mx + c$. The relevant example for our purposes is the case when the manifold Σ represents the pulsar parameters $\vec{\xi} = (f_{(0)}, \{f_{(n)}\}, \mathbf{n})$ and M is the time-frequency plane. The pattern in M is described by the Doppler shift formula of equation (2.4). Each value of $\vec{\xi}$ determines the frequency evolution $f(t)$ and thus determines a curve in the time-frequency plane.

Given a set of observations $\{x_i\}$ with each x_i belonging to M, we ask if there is an underlying pattern describing these points and whether this pattern is described by a hypersurface belonging to $\mathcal{C}$. Consider first the idealized case when there is no noise and the points $\{x_i\}$ actually do follow the pattern and lie on one single hypersurface belonging to $\mathcal{C}$ corresponding to the parameter value $\hat{\mu} \in \Sigma$. How would we go about finding $\hat{\mu}$ if all we are given is the collection $\{x_i\}$? For every x_i, the idea is to first find the set of points $\mathcal{U}_i$ in parameter space consistent with x_i; the true parameter value $\hat{\mu}$ must certainly lie within this set. In the straight line example, all the lines passing through the observed point would be consistent with the observation. Repeating this for every observation x_i, we obtain a collection of subsets $\{\mathcal{U}_i\}$. The true parameter value $\hat{\mu}$ must lie in each $\mathcal{U}_i$ and therefore it must also lie in the intersection

$$\hat{\mu} \in \bigcap_i \mathcal{U}_i \,. \qquad (3.1)$$

See figure 3. If k is the dimensionality of Σ, then we need at least k different x_i's in order to ensure that $\hat{\mu}$ can be found uniquely. Thus, in this idealized case where noise is not present, we would need only two observations to detect a straight line. Similarly for the pulsar case when equation (2.4) is the master equation and if we are searching for s spindown parameters, we would need only $3+s$ observations to determine the pulsar parameters. This is, of course, not true when noise is present.

In realistic situations, the presence of noise will ensure that, in general, there is no point which is consistent with *all* the x_i's, in other words, $\cap_i \mathcal{U}_i$ is the empty set. In this case we proceed as follows: to each $\mu \in \Sigma$, assign an integer $n(\mu)$ (the *number count*), which is equal to the number of $\mathcal{U}_i$'s which contain μ. The result is then a histogram in parameter space. This procedure which maps us from a set of observations to a histogram in parameter space will be called the Hough transform. The best candidate for the true parameter $\hat{\mu}$ is then the point at which the number count is maximum. Alternatively, we could set an appropriate threshold $n_{\rm th}$ on the number count and select all points in Σ at which the number count exceeds $n_{\rm th}$. These selected parameter space points would be candidates for a possible detection and, if we are performing a multi stage hierarchical search, would be further analyzed in the next step.

In real experiments, we cannot perform a parameter space search with infinite resolution. Therefore we need to consider the discrete case when we have a finite resolution for the observations and also a grid on parameter space. In this case, observations correspond to *pixels* in M. The general procedure is essentially the same as in the discrete case and is depicted schematically in figure 4: we look for pixels in parameter space which are consistent with the observations. There is, however, one technical

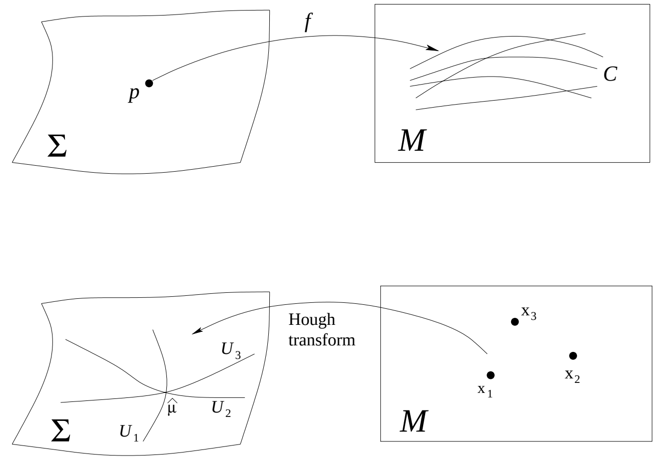

FIG. 3: A schematic depiction of the Hough transform in the absence of noise. The top figure shows the parameter space Σ and the space of observations M. The space of expected patterns in M is a set $\mathcal{C}$ of hypersurfaces in M. The function $f : \Sigma \to \mathcal{C}$ provides a one-one correspondence between Σ and $\mathcal{C}$. The lower figure shows the Hough transform itself: Every observation x_i is mapped via the Hough transform into a hypersurface $\mathcal{U}_i$ in parameter space which is consistent with the observation. The intersection of all the $\mathcal{U}_i$'s contains the true source parameter $\hat{\mu}$.

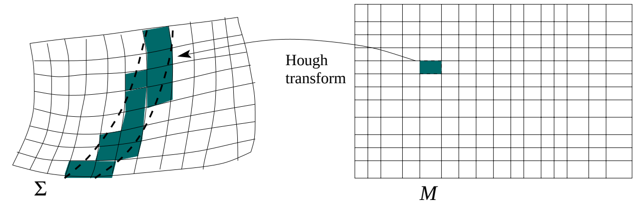

FIG. 4: A schematic view of the Hough transform for the discrete case. An observation consists of a pixel in M which goes over to the region enclosed between the dotted lines under the Hough transform. This in turn leads to a selection of pixels in parameter space. The shaded pixels are the ones which get selected and are the ones consistent with the observation.

difference namely, since each observation is an extended region in M, the points in parameter space consistent with this observation do not constitute sharp hypersurface $\mathcal{U}_i$, but instead give a region $\tilde{\mathcal{U}}_i$ bounded by two such hypersurfaces. Given such a region, we can then select pixels in parameter space. Since a pixel in parameter space might intersect more that one $\tilde{\mathcal{U}}_i$, we need an unambiguous criteria to select pixels in parameter space in order to ensure that each pixel gets selected at most once by an observation. Given such a criterion, we can continue the earlier strategy and construct a histogram in parameter space by assigning a number count to each pixel in parameter space. The pixel with the largest number count is then our best candidate for a detection.

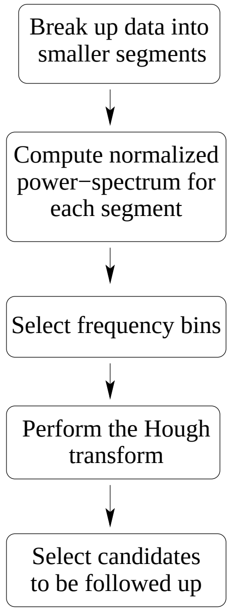

FIG. 5: A single stage of a hierarchical CW search involving the Hough transform. The starting point is to break up the data with total observation time T_{obs} into N segments and to compute the Fourier transform of each segment. The next step is to select frequency bins from each SFT by setting a threshold on the normalized power spectrum and use the selected frequency bins to construct a Hough map. The output is then a set of candidates in parameter space obtained by setting a threshold on the Hough number count. Alternatively, one could also set upper limits by using the pixel with the maximum number count.

IV. THE HOUGH TRANSFORM WITH NON-DEMODULATED DATA

The steps involved in a single incoherent stage of the search are outlined in figure 5. In this search, one starts by breaking up the input data of duration $T_{\rm obs}$ into N segments each with a duration of $T_{\rm coh}$ which would be equal to $T_{\rm obs}/N$ if there were no gaps in the data. Then compute the Fourier transform of each data segment to obtain N SFTs. Select frequency bins in each SFT by setting a threshold on the normalized power spectrum. This produces a distribution of points in the time-frequency plane — the manifold M — most of which are noise but some excess of which are hopefully present along one or more signal patterns given by equation (2.4). Having selected points in the time-frequency plane, go through the Hough transform algorithm to obtain the Hough map, i.e. the histogram, in parameter space Σ. The details follow.

A. Notation and conventions

As described above, the first step in the search is to break up the available data into N segments. Successive data segments need not be contiguous in time; there may be gaps in the data stream representing times when the detector is not in lock or when the data is not reliable. We assume that the different data segments have the same time duration. Label the different segments by $a = 0, 1 \ldots (N-1)$ and denote the start time of each segment by t_a which will often be called the *timestamp* of the a^{th} data segment. Let each segment consist of M data points and let the time baseline of each segment be $T_{\rm coh}$.

Let us now focus on the a^{th} data segment which covers the time interval $[t_a, t_a + T_{\rm coh}]$. Let $x(t)$ be the detector output which is sampled at times $t_n = t_a + n\Delta t$ with $n = 0, 1, \ldots (M-1)$. Here the data segment has been subdivided into M sub-segments with the times t_n defined to be at the start of each sub-segment; this implies $\Delta t = T_{\rm coh}/M$. Denote the sequence of data points thus obtained by $\{x_n\}$ where $x_n \equiv x(t_n)$.

Our convention for the Discrete Fourier Transform (DFT) of $\{x_n\}$ is

$$\tilde{x}_k = \Delta t \sum_{j=0}^{M-1} x_j e^{-2\pi i jk/M} \tag{4.1}$$

where $k = 0, 1 \ldots (M-1)$. For $0 \leq k \leq \lfloor M/2 \rfloor$, the frequency index k corresponds to a physical frequency of $f_k = k/T_{\rm coh}$ with $\lfloor . \rfloor$ denoting the integer part of a given real number. The values $\lfloor M/2 \rfloor < k \leq M-1$ correspond to negative frequencies given by $f_k = (k-M)/T_{\rm coh}$.

The detector output $x(t)$ at any time t is the sum of noise $n(t)$ and a possible gravitational wave signal $h(t)$ of known form:

$$x(t) = n(t) + h(t) \,. \tag{4.2}$$

In the remainder of this paper, unless otherwise stated, the stochastic process $n(t)$ is assumed to be stationary and Gaussian with zero mean.

In the continuous case, when the observation time is infinite, the single-sided power spectral density (PSD) $S_n(f)$ for $f \geq 0$ is defined as the Fourier transform of the auto-correlation function:

$$S_n(f) = 2\int_{-\infty}^{\infty} \langle n(t) n(0) \rangle e^{-2\pi i f t} dt \tag{4.3}$$

where $\langle \cdot \rangle$ denotes the ensemble average.

The *normalized power* is a dimensionless quantity defined as

$$\rho_k = \frac{|\tilde{x}_k|^2}{\langle |\tilde{n}_k|^2 \rangle} \tag{4.4}$$

It can be shown that $\langle |\tilde{n}_k|^2 \rangle$ is related to the PSD:

$$\langle |\tilde{n}_k|^2 \rangle \approx \frac{M\Delta t}{2} S_n(f_k) = \frac{T_{\rm coh}}{2} S_n(f_k) \,. \tag{4.5}$$

Thus:

$$\rho_k \approx \frac{2|\tilde{x}_k|^2}{T_{\rm coh} S_n(f_k)} \,. \tag{4.6}$$

Naturally, the PSD must be estimated in a way that is not biased by any signal power that may be present.

B. Implementation

The implementation choices we present here mostly correspond to those that have been implemented in the Hough analysis code which is publicly available as part of the LIGO Algorithms Library (LAL) [12], and will be used to analyze the data from the GEO and LIGO detectors.

Restriction on T_{coh}: For non-demodulated data, the coherent integration time $T_{\rm coh}$, i.e. the timebaseline of the SFTs, cannot be arbitrarily large. This restriction comes about because we would like the signal power to be concentrated in half a frequency bin but the signal frequency is changing in time due to the Doppler modulation and also due to the spindown of the star. If $\dot{f}$ is the time-derivative of the signal frequency at any given time, in order for the signal not to shift by more than half a frequency bin, we must have $|\dot{f}|T_{\rm coh} < (2T_{\rm coh})^{-1}$, i.e.

$$T_{\rm coh} < \sqrt{\frac{1}{2|\dot{f}|_{\rm max}}} \tag{4.7}$$

where by $|\dot{f}|_{\rm max}$ we mean the maximum possible value of $|\dot{f}|$ for all allowed values of the shape parameters $\vec{\xi}$. The time variation of $f(t)$ is given by equation (2.4) and is due to two effects: the spindown of the star, and the Doppler modulation due to the Earth's motion. We shall assume that the Doppler modulation is the dominant effect [25]. Thus we can estimate $\dot{f}$ by keeping $\hat{f}$ fixed and differentiating $\mathbf{v}(t)$ in equation (2.4):

$$\dot{f} \approx \frac{\hat{f}}{c}\frac{d\mathbf{v}}{dt}\cdot\mathbf{n} \leq \frac{\hat{f}}{c}\left|\frac{d\mathbf{v}}{dt}\right| . \tag{4.8}$$

The important contribution to the acceleration $d\mathbf{v}/dt$ is from the daily rotation of the Earth:

$$|\dot{f}|_{\rm max} = \frac{\hat{f}}{c}\cdot\frac{v_e^2}{R_e} = \frac{\hat{f}}{c}\cdot\frac{4\pi^2 R_e}{T_e^2} \tag{4.9}$$

where v_e is the magnitude of the velocity of Earth around its axis, T_e the length of a day and R_e the radius of Earth. Substituting numerical values we get

$$T_{\rm coh} < 50\,{\rm min}\times\sqrt{\frac{500\,{\rm Hz}}{\hat{f}}} . \tag{4.10}$$

In this paper, we shall mostly use $T_{\rm coh} = 30$min as the canonical reference value.

Selecting frequency bins: Having computed the normalized power, our next step now is to select frequency bins from the SFT. The simplest method of selecting frequency bins is to set a threshold $\rho_{\rm th}$ on ρ_k; i.e. we select the k^{th} frequency bin if $\rho_k \geq \rho_{\rm th}$ and reject it otherwise. Alternatively [19, 21], we could impose additional conditions such as requiring that $\rho_k > \rho_{k+1}$ and $\rho_k > \rho_{k-1}$, i.e. the k^{th} bin is selected if ρ_k exceeds the threshold and is, in addition, a *local maxima*. This can be extended further by including more than just the two neighboring frequency bins. While it is relatively easy to investigate these alternate strategies for non-demodulated data, the analysis becomes more complicated for demodulated data. Furthermore, while these alternate methods might be more robust against spectral disturbances, the analysis of the statistics follows the same general scheme and the results are not qualitatively different. Thus, for the purposes of this paper, we will describe only the simple thresholding scheme for selecting frequency bins. The optimal choice of the threshold $\rho_{\rm th}$ will be described later in section V B.

Solving the master equation: As discussed in the previous section, to perform the Hough transform, we must find all the points in parameter space which are consistent with a given observation. In this case, the observation is a frequency f_k selected using a threshold $\rho_{\rm th}$ in say, the a^{th} SFT corresponding to a timestamp t_a. This corresponds to a frequency bin $(f_k - \frac{1}{2}\delta f, f_k + \frac{1}{2}\delta f)$ where $\delta f = T_{\rm coh}^{-1}$ is the frequency resolution of the SFT. The parameters $\vec{\xi}$ of the signal are the frequency, spindown parameters, and the sky-positions: $\vec{\xi} = (f_{(0)}, \{f_{(n)}\}, \mathbf{n})$. Corresponding to $(f_k - \frac{1}{2}\delta f, f_k + \frac{1}{2}\delta f)$, we must then find all the possible values of $\vec{\xi}$ which satisfy the master equation (2.4).

To understand this better, let us first fix the values of the frequency $f_{(0)}$ and the spindown parameters $\{f_{(n)}\}$ so that $\hat{f}(t)$ is also fixed. Ignore, for the moment, the frequency resolution δf. From equation (2.4), we see that all the values of $\mathbf{n}$ consistent with the observation $f(t)$ must satisfy

$$\cos\phi = \frac{\mathbf{v}(t)\cdot\mathbf{n}}{v(t)} = \frac{c}{v(t)}\frac{f(t)-\hat{f}(t)}{\hat{f}(t)} \tag{4.11}$$

where ϕ is the angle between $\mathbf{v}(t)$ and $\mathbf{n}$. This implies that the angle ϕ must be constant; in other words, the set of sky positions consistent with an observation $f(t)$ form a *circle* in the celestial sphere centered on the vector $\mathbf{v}$ (see figure 6) [26]. If the frequency $f(t)$ is smeared over a frequency bin $(f_k - \frac{1}{2}\delta f, f_k + \frac{1}{2}\delta f)$, the set of points consistent with an observation must correspond to an *annulus* the width $\delta\phi$ of which is easily calculated using equation (4.11):

$$\delta\phi \approx \frac{c}{v}\frac{\delta f}{\hat{f}\sin\phi} . \tag{4.12}$$

The annuli are very thick at points where $\sin\phi$ is small, i.e. when $\mathbf{n}$ is almost parallel or anti-parallel to $\mathbf{v}(t)$ and

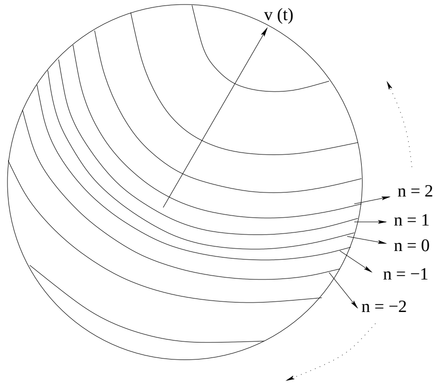

FIG. 6: The set of sky positions consistent with a given frequency bin at a given time correspond to annuli on the celestial sphere. These annuli are centered on the velocity vector $\mathbf{v}$, they are thin when perpendicular to $\mathbf{v}$ and thick when nearly parallel. The circle with $n = 0$ corresponds to $f = \hat{f}$.

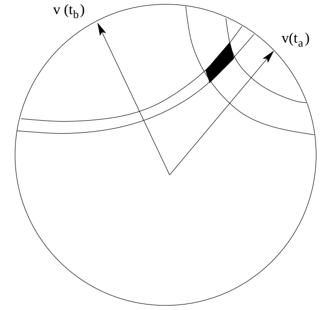

FIG. 7: Two intersecting annuli. The two timestamps t_a and t_b are sufficiently different from each other so that the velocities $\bar{\mathbf{v}}(t_a)$ and $\bar{\mathbf{v}}(t_b)$ are not parallel to each other. This causes the annuli constructed at different timestamps to intersect. The shaded region is the intersection and there is a corresponding region (not shown) on the far side of the sphere.

very thin when perpendicular. This is depicted schematically in figure 6. The circles on the celestial sphere are labelled by an integer n such that the frequency $f = \hat{f} + n\delta f$ corresponds to the angle ϕ_n given by

$$\cos\phi_n = \frac{nc\delta f}{v\hat{f}} . \tag{4.13}$$

The lower limit on the width of the annuli is provided by setting $\phi = \pi/2$ in equation (4.12):

$$(\delta\phi)_{\min} = \frac{c}{v}\frac{\delta f}{\hat{f}} = \frac{c}{v\hat{f}T_{\mathrm{coh}}} \tag{4.14}$$
$$= 4.8\times 10^{-3}\,\mathrm{rad}\times\left(\frac{1\mathrm{hr}}{T_{\mathrm{coh}}}\right)\left(\frac{500\,\mathrm{Hz}}{\hat{f}}\right)\left(\frac{10^{-4}}{v/c}\right) .$$

The upper limit on the annuli width $(\delta\phi)_{\max}$ is found by setting $\sin\phi \approx \phi \approx (\delta\phi)_{\max}$ which gives

$$(\delta\phi)_{\max} = \sqrt{\frac{\delta f}{\hat{f}}\frac{c}{v}} = \sqrt{\frac{c}{v\hat{f}T_{\mathrm{coh}}}} \tag{4.15}$$
$$= 7.3\times 10^{-2}\,\mathrm{rad}\times\left(\frac{1\mathrm{hr}}{T_{\mathrm{coh}}}\right)^{\frac{1}{2}}\left(\frac{500\mathrm{Hz}}{\hat{f}}\right)^{\frac{1}{2}}\left(\frac{10^{-4}}{v/c}\right)^{\frac{1}{2}} .$$

Therefore, the thick annuli are about 10 times thicker than the thin ones. Different frequency bins selected at the same time will correspond to non-intersecting annuli as shown in figure 6. However, for frequency bins selected from SFTs at different time stamps, say t_a and t_b, the annuli will usually intersect because the velocity vectors $\bar{\mathbf{v}}(t_a)$ and $\bar{\mathbf{v}}(t_b)$ will not, in general, be parallel to each other; see figure 7.

Resolution in the space of sky-positions: In order to search for pulsar signals in a given portion of the sky, we must choose a tiling for the sky patch. Given the calculation of the annuli width above, we choose the pixel size $\delta\theta$ of the grid to be some fraction, say at most half, of the width $(\delta\phi)_{\min}$ of the thinnest annulus. While this educated guess for the pixel size is sufficient for the purposes of this paper, the correct choice of pixel size in the sky patch, and also in the entire parameter space, should use the parameter space metric introduced in [22]. The analysis of this metric for the Hough search will be presented elsewhere, and for now we shall simply use $\delta\theta = \frac{1}{2}(\delta\phi)_{\min}$.

There are, of course, many ways of placing the templates on the space of sky locations. We could, for instance, choose a uniform grid on $(\cos\theta,\phi)$ where θ is the polar angle and ϕ the azimuthal angle. Alternatively, instead of using the celestial sphere directly, we could choose a rectangular grid on the plane and map it to the sphere using the stereographic projection. Under the stereographic map, circles on the sphere are mapped to circles on the plane. The distortion of the rectangular pixels produced by the stereographic map will not be too severe as long as the sky patch is not too large. To perform an all sky search, we will then have to break up the celestial sphere into several patches and perform a different stereographic projection on each patch.

Having selected an annulus and having chosen a tiling on our sky-patch, we now need a criteria for selecting a pixel if it intersects an annulus. Our criteria is to select a pixel if its center lies within an annulus. Under such a criteria, a given pixel can then be selected by at most one annulus and the pixels selected by all the annuli together will exactly cover the sphere.

Resolution in the space of spindown parameters: In the absence of a proper analysis of the parameter space metric which determines the correlations between the different spindown parameters, we shall just use the obvious estimate for the resolution $\delta f_{(n)}$ in the space of spindown parameters:

$$\delta f_{(n)} = n!\frac{\delta f}{T^n_{\mathrm{obs}}} . \tag{4.16}$$

As an example, for the first spindown parameter:

$$\delta f_{(1)} = (3.1\times 10^{-10}\text{Hz/s})\times\frac{1000}{N}\left(\frac{1800\text{s}}{T_{\text{coh}}}\right)^2 . \quad (4.17)$$

We now need to choose the range of values $-f_{(n)}^{\max} < f_{(n)} < f_{(n)}^{\max}$ and the largest number of spindown parameters $s_{\max}$ to be searched over. Assuming that the pulsar's frequency evolution is well-represented by a Taylor expansion, we get

$$f_{(n)}^{\max} = n!\frac{\hat{f}_{\max}}{\tau^n} \quad (4.18)$$

where τ is the age of the pulsar and $\hat{f}_{\max}$ is the largest intrinsic frequency that we search over. We include the n^{th} spindown parameter only if the resolution defined by equation (4.16) is not too coarse compared to $f_{(n)}^{\max}$:

$$\delta f_{(n)} < f_{(n)}^{\max} . \quad (4.19)$$

Since $T_{\text{obs}} \ll \tau$, $f_{(n)}^{\max}$ decreases with increasing n much faster than $\delta f_{(n)}$; this implies that there must exist a value $s_{\max}$ such that equation (4.19) is satisfied for all $n \leq s_{\max}$ and is violated for all $n > s_{\max}$. Any spindown parameter of order greater than $s_{\max}$ does not significantly affect the result of the Hough transform. As an example, if we wish to search for pulsars whose age is at least $\tau = 40$yrs, then for $\hat{f}_{\max} = 1000$Hz, it is easy to check that we get $s_{\max} = 3$. In other words, to look for pulsars which are as young as 40yr, we must include at least 3 spindown parameters in our search.

On the other hand, in some cases, computational requirements might dictate that we can only search over, say, one spindown parameter. This automatically sets a lower limit on the age of the pulsar that we can search over because then the second spindown parameter must satisfy $\delta f_{(2)} > f_{(2)}^{\max}$ which leads to

$$\tau > 216\text{yr}\times\frac{N}{2000}\left(\frac{\hat{f}_{\max}}{1000\text{Hz}}\right)^{1/2}\left(\frac{T_{\text{coh}}}{1800\text{s}}\right)^{3/2} . \quad (4.20)$$

Finally, the finite length of T_{coh} itself leads to a lower bound on τ. If $f_{(n)}$ is too large, then the signal power may no longer be concentrated in a single frequency bin and the assumption of neglecting spindown parameters which was used to derive equation (4.10) will no longer be valid. To be certain that the spindown will not cause the signal to move by more than half a frequency bin, we must have $f_{(n)}^{\max}T_{\text{coh}}^n < n!\delta f/2$ which implies

$$\tau > \left(\frac{2\hat{f}_{\max}T_{\text{coh}}^{n+1}}{n!}\right)^{1/n} . \quad (4.21)$$

The most stringent limit is obtained for $n = 1$:

$$\tau > 103\text{yr}\times\frac{\hat{f}_{\max}}{1000\text{Hz}}\left(\frac{T_{\text{coh}}}{1800\text{s}}\right)^2 . \quad (4.22)$$

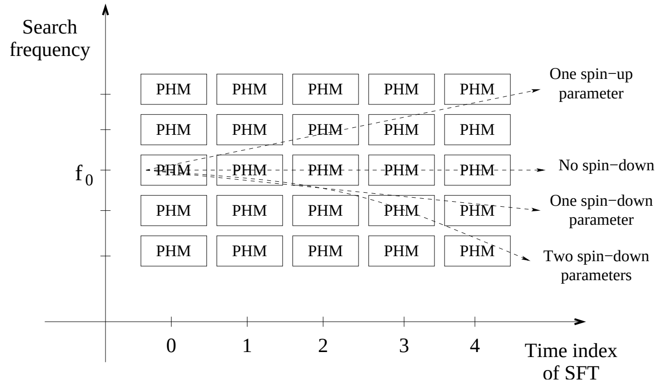

FIG. 8: A Partial Hough Map (PHM) is a Hough map, i.e. a histogram in the (α,δ) plane, constructed from all the frequencies selected at a given time and for a given value of the instantaneous frequency $\hat{f}_0$. A total Hough map is obtained by summing over the appropriate Hough maps. The PHMs to be summed over are determined by the choice of spindown parameters which give a *trajectory* in the time-frequency plane. For example, a single spin down parameter will give a straight line as shown is the figure while two spindown parameters will lead to a parabola.

Note that this restriction will not be present if we use demodulated data as input for the Hough transform.

Partial and total Hough maps: As described above, for a given frequency bin selected at a given time-stamp and for a given value of the instantaneous frequency $\hat{f}$, we can find the set of sky locations which are consistent with the master equation (2.4). In other words, every pixel in the sky-patch either gets selected or rejected and this gives a histogram in the (α,δ) plane consisting of ones or zeros; α and δ are coordinates on the sky-patch. Such a collection of ones and zeros on the sky-patch is called a *Partial Hough Map* (PHM). The number of PHMs required at any given time depends on the frequency band Δf_b that one is searching over and is simply given by the number $\Delta f_b/\delta f = T_{\text{obs}}\Delta f_b$.

Given a set of PHM's for every time interval, and given a set of spindown parameters that one wishes to search for, the *Total Hough Map* (THM) is obtained by summing the appropriate partial Hough maps. To see how this comes about, consider the case when we are searching for some spindown parameter $\{f_{(n)}\}$ with $n = 1, 2, \cdots$. The instantaneous frequency changes with time according to equation (2.5) This can be viewed as a trajectory in the time-frequency plane. A single spindown parameter will give a straight line, two spindown parameters a parabola and so on. Thus for each time-stamp t_a, we can find the appropriate PHM by looking at which frequency bin this trajectory intersects (see figure 8). For a given choice of spindown parameters, the THM is obtained by summing over the appropriate PHMs. Repeating this for every set of frequency and spindown parameters we wish to search over, we obtain a number of THMs and the collection of all these THMs represent our final histogram in parameter space.

Look up tables: The procedure described thus far is, in principle, enough to produce a complete Hough map in parameter space. However, it is possible to enormously reduce the computational cost by using *Look Up Tables* (LUTs) which we now describe. Assume that we have managed to find all the annuli for a given time-stamp t_a and for a given search frequency $\hat{f}$. To construct the PHM for t_a and $\hat{f}$, we just need to select the appropriate annuli out of all the ones that we have found. Very importantly, it turns out that in most cases, the annuli are relatively insensitive to changes in $\hat{f}$ and can therefore be re-used a large number of times.

To see this, look at how the solutions of equation (2.4) depend on the search frequency $\hat{f}$. We want to calculate the maximum number κ of frequency bins that $\hat{f}$ can be changed by so that the annuli change by only a fraction r of the quantity $(\delta\phi)_{\rm min}$ defined in equation (4.15). As discussed earlier, if we restrict ourselves to discrete frequencies, the annuli corresponding to every given value of $\hat{f}$ are parameterized by an integer n according to equation (4.13). For a fixed value of n, by how much does ϕ_n change when $\hat{f}$ is varied? To answer this, differentiate equation (4.13) with respect to $\hat{f}$:

$$\frac{d\hat{f}}{\hat{f}^2} = \frac{1}{n\delta f}\frac{v}{c}\sin\phi_n\, d\phi_n = \frac{\tan\phi_n}{\hat{f}} d\phi_n\,. \tag{4.23}$$

Set $d\phi_n = r(\delta\phi)_{\rm min}$ and $d\hat{f} = \kappa\delta f = \kappa/T_{\rm coh}$ to obtain

$$\kappa = \frac{rc}{v}\tan\phi_n = \frac{rc}{v}\sqrt{\frac{n_0^2}{n^2} - 1} \tag{4.24}$$

where $n_0 = v\hat{f}/(c\delta f)$. Consider separately the two regimes when $\phi_n \sim \pi/2$ (i.e. $n \sim 0$) and $\phi_n \sim 0, \pi$ (i.e. $n \sim \pm n_0$). When $\phi_n = \pi/2$, then κ is infinite which indicates that a LUT is excellent in this regime. On the other hand, $\kappa = 0$ for $\phi_n = 0, \pi$. However, since the resolution in ϕ is finite, instead of $\phi_n = 0$, it is more appropriate to take the worst case scenario as $\phi_n = (\delta\phi)_{\rm min}$ so that

$$\kappa \approx \frac{rc}{v}(\delta\phi)_{\rm min} = 40r\left(\frac{500\,{\rm Hz}}{\hat{f}}\right)^{\frac{1}{2}}\,. \tag{4.25}$$

Thus, in this worst case scenario, for a frequency of 500 Hz and a tolerance of $r = 0.1$, the LUT will be valid for 4 frequency bins. Furthermore, due to the presence of the function $\tan\phi$ in equation (4.24), κ increases rapidly with increasing ϕ_n (i.e. decreasing n). As an example, take $T_{\rm coh} = 1800$s, $\hat{f} = 500$ Hz, and $v/c = 10^{-4}$ so that $n_0 = 90$. Then, even for $n = 89$, we get $\kappa = 1500r$; thus with say $r = 0.1$, the LUT is valid for about 150 frequency bins. This shows that generically, the use of LUTs leads to very large savings in computational costs.

V. STATISTICAL PROPERTIES OF THE HOUGH MAPS

This section is divided into three parts: The probability distribution of the number counts is calculated in section V A, section V B optimizes the various thresholds and section V C estimates the sensitivity of the Hough search.

A. The number count distribution

The frequency bins that are fed into the Hough transform are the ones such that their normalized power ρ_k defined in equation (4.4) exceeds a threshold $\rho_{\rm th}$. Assuming that the noise is stationary, has zero mean, and is Gaussian, from equation (4.4), we get

$$2\rho_k = z_1^2 + z_2^2 \tag{5.1}$$

where

$$z_1 = \frac{\sqrt{2}{\rm Re}[\tilde{x}_k]}{\sqrt{\langle|\tilde{n}_k|^2\rangle}} \qquad \text{and} \qquad z_2 = \frac{\sqrt{2}{\rm Im}[\tilde{x}_k]}{\sqrt{\langle|\tilde{n}_k|^2\rangle}}\,. \tag{5.2}$$

Assuming that ${\rm Re}[\tilde{n}_k]$ and ${\rm Im}[\tilde{n}_k]$ are independent random variables with equal variance, it is easy to show that their variance must be equal to $\langle|\tilde{n}_k|^2\rangle/2$. Therefore, the random variables z_1 and z_2 are normally distributed and have unit variance. Thus $2\rho_k$ must be distributed according to a non-central χ^2 distribution with 2 degrees of freedom. The non centrality parameter λ_k can be easily calculated:

$$\lambda_k = (\mathbf{E}[z_1])^2 + (\mathbf{E}[z_2])^2 = \frac{4|\tilde{h}(f_k)|^2}{T_{\rm coh}S_n(f_k)}\,. \tag{5.3}$$

Thus the distribution of ρ_k is

$$\begin{aligned} p\,(\rho_k|\lambda_k) &= 2\chi^2(2\rho_k|\,2,\lambda_k) \\ &= \exp\left(-\rho_k - \frac{\lambda_k}{2}\right) I_0(\sqrt{2\lambda_k\rho_k}) \end{aligned} \tag{5.4}$$

where I_0 is the modified Bessel's function of zeroth order. As expected, $p\,(\rho_k|\lambda_k)$ reduces to the exponential distribution in the absence of a signal when $\lambda = 0$.

The mean and variance for this distribution are respectively

$$\mathbf{E}[\rho_k] = 1 + \frac{\lambda_k}{2} \qquad \text{and} \qquad \sigma^2[\rho_k] = 1 + \lambda_k\,. \tag{5.5}$$

The probability η that a given frequency bin is selected is

$$\eta(\rho_{\rm th}|\lambda) = \int_{\rho_{\rm th}}^{\infty} p(\rho|\lambda)\,{\rm d}\rho \tag{5.6}$$

where we have dropped the subscript k for notational simplicity; it is understood that ρ and λ always refer to

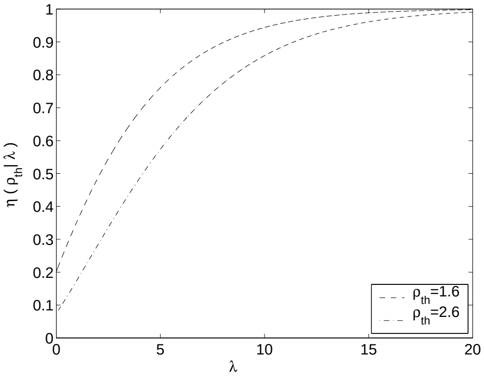

FIG. 9: Plot the detection probability $\eta(\rho_{\rm th}|\lambda)$ as a function of λ for $\rho_{\rm th} = 1.6$ and 2.6.

one of the Fourier frequency bins. The false alarm and false dismissal probabilities for the frequency bin selection are defined respectively as

$$\alpha(\rho_{\rm th}) = \int_{\rho_{\rm th}}^{\infty} p(\rho|0)\,\mathrm{d}\rho = e^{-\rho_{\rm th}}\,, \qquad (5.7)$$

$$\beta(\rho_{\rm th}|\lambda) = 1 - \eta(\rho_{\rm th}|\lambda) = \int_0^{\rho_{\rm th}} p(\rho|\lambda)\,\mathrm{d}\rho\,. \qquad (5.8)$$

Clearly, $\eta = \alpha$ when no signal is present and η becomes larger when the signal becomes stronger and $\eta \to 1$ when $\lambda \to \infty$. Figure 9 shows $\eta(\rho_{\rm th}|\lambda)$ as a function of the non-centrality parameter λ for two different values of $\rho_{\rm th}$. For small λ:

$$\eta(\rho_{\rm th}|\lambda) = \alpha\left\{1 + \frac{\rho_{\rm th}}{2}\lambda + \mathcal{O}(\lambda^2)\right\}\,. \qquad (5.9)$$

This expansion will be very useful when we restrict ourselves to the case of small signals.

In the presence of a signal, the non-centrality parameter λ_k is *not* constant across different SFTs. The reason for this is two-fold: First, the noise may be significantly non-stationary. Secondly, and more fundamentally, the observed signal power $|\tilde{h}|^2$ changes because of the amplitude modulation of the signal caused by the non-uniform antenna pattern of the detector. Therefore, the detection probability η changes across SFTs. In what follows, we shall neglect this effect and take λ and η to be constant for different SFTs.

Under this assumption, the probability of measuring a number count n in a pixel of a Hough map constructed from N SFTs is

$$p(n|\rho_{\rm th},\lambda) = \binom{N}{n}\eta^n(1-\eta)^{N-n}\,. \qquad (5.10)$$

The mean and variance of the number count are respectively

$$\bar{n} = N\eta \qquad \text{and} \qquad \sigma^2 = N\eta(1-\eta)\,. \qquad (5.11)$$

In the absence of a signal, $\eta = \alpha$ so that

$$p(n|\rho_{\rm th},0) = \binom{N}{n}\alpha^n(1-\alpha)^{N-n}\,. \qquad (5.12)$$

Candidates for detection or for further analysis are selected by setting a threshold $n_{\rm th}$ on the number count. Based on this, we can define the false alarm and false dismissal rates respectively as:

$$\alpha_H(n_{\rm th},\rho_{\rm th},N) = \sum_{n=n_{\rm th}}^{N} p(n|\rho_{\rm th},0)\,, \qquad (5.13)$$

$$\beta_H(n_{\rm th},\rho_{\rm th},\lambda,N) = \sum_{n=0}^{n_{\rm th}-1} p(n|\rho_{\rm th},\lambda)\,. \qquad (5.14)$$

These two quantities determine the significance and the sensitivity of the Hough search and will play an important role in the rest of this paper.

B. Optimal choice of the thresholds

In order to carry out the Hough search, we have to set two thresholds: the threshold $\rho_{\rm th}$ on the normalized power and the threshold $n_{\rm th}$ on the number count.

The value of $n_{\rm th}$ is determined by the false alarm rate $\alpha_H^\star$ that depends on the number of candidates that we can feasibly follow up.

The value of $\rho_{\rm th}$ is chosen in such as way so as to make the search as powerful as possible. We present two criteria that yield the same result for small signals and for large N.

Maximizing the critical ratio: For the Hough number count, we can define a random variable called the *critical ratio* as follows

$$\Psi(\eta,\alpha) = \frac{n - N\alpha}{\sqrt{N\alpha(1-\alpha)}}\,, \qquad (5.15)$$

This quantity is a measure of the "significance" of a measured value n with respect to the expected value $N\alpha$ in the absence of any signal, weighted by the expected fluctuations of the noise. In the presence of a signal, the expected value of the critical ratio is

$$\bar{\Psi}(\eta,\alpha) = \frac{N\eta - N\alpha}{\sqrt{N\alpha(1-\alpha)}}\,, \qquad (5.16)$$

Recall that η and α depend on the threshold $\rho_{\rm th}$ according to equations (5.7) and (5.8) respectively. Thus, our criterion for choosing the threshold is to maximize $\bar{\Psi}(\eta,\alpha)$ with respect to $\rho_{\rm th}$. In the case of small signals where $\eta \approx \alpha(1+\rho_{\rm th}\lambda/2)$, the condition

$$\frac{\partial\bar{\Psi}}{\partial\rho_{\rm th}} = 0 \qquad (5.17)$$

leads to

$$\ln\alpha = 2(\alpha - 1) \qquad (5.18)$$

which yields $\rho_{\rm th} = 1.6$ or equivalently, $\alpha = 0.20$.

The Neyman-Pearson criterion: An alternative method of choosing $\rho_{\rm th}$ is based on the Neyman-Pearson criterion which tells us to minimize the false dismissal rate β_H for a given value $\alpha_H^\star$ of the false alarm rate. For weak signals, this requirement uniquely determines $\rho_{\rm th}$ and, as we shall see, this agrees with the previous criterion.

In practice, taking N and λ to be fixed parameters, this is the procedure:

i. First choose a value $\alpha_H^\star$ for the largest false alarm rate α_H that we can allow.

ii. Invert the equation $\alpha_H(\rho_{\rm th}, N, n_{\rm th}) \leq \alpha_H^\star$ to obtain $n_{\rm th}(\rho_{\rm th}, N, \alpha_H^\star)$.

iii. Substitute the value of $n_{\rm th}$ thus obtained in the expression for the false dismissal $\beta_H(n_{\rm th}, \rho_{\rm th}, \lambda, N)$. This gives β_H as a function of $(\rho_{\rm th}, \lambda, N, \alpha_H^\star)$.

iv. Minimize β_H as a function of $\rho_{\rm th}$. Let $\rho_{\rm th}^\star$ be the value that minimizes β_H.

v. Using $n_{\rm th}(\rho_{\rm th}, N, \alpha_H^\star)$ derived in step (ii) above, obtain $n_{\rm th}^\star = n_{\rm th}(\rho_{\rm th}^\star, N, \alpha_H^\star)$.

This procedure is also applicable if we choose a different method of selecting frequency bins other than simple thresholding, such as, for example the peak selection criteria mentioned towards the end of section IV A.

The results of the optimization procedure described above are shown in figures 10, 11 and 12. Figure 10 shows the value of the number count threshold $n_{\rm th}$ obtained as described in step (ii). In this figure, instead of $\rho_{\rm th}$, we have chosen the false alarm rate $\alpha = e^{-\rho_{\rm th}}$ as the independent variable; α is the false alarm rate for selecting frequency bins and is not to be confused with α_H. Figure 10 also shows an analytic approximation to $n_{\rm th}$ obtained below in equation (5.22). Using this result for $n_{\rm th}$, figure 11 shows β_H as a function of $\alpha = e^{-\rho_{\rm th}}$. The optimal choice $\rho_{\rm th}^\star$ of $\rho_{\rm th}$ is when β_H is a minimum and, for small signals, this happens at approximately $\rho_{\rm th}^\star = 1.6$ which corresponds to $\alpha^\star := e^{-\rho_{\rm th}^\star} = 0.20$. Finally, figure 12 shows the minimum value of β_H obtained by this optimization as a function of the signal strength λ and for two different values of N.

The Gaussian approximation: To better understand the statistics, it is useful to carry out the above steps analytically by taking n to be a continuous variable and by approximating the binomial distribution by a Gaussian with the appropriate mean and variance:

$$p(n|\rho_{\rm th}, \lambda) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(n-N\eta)^2/2\sigma^2} \,. \tag{5.19}$$

This is a very good approximation when N is large and η is not too close to 0 or 1. If n is chosen to lie within $[0, N]$, the distribution is properly normalized only approximately. For simplicity, in what follows we shall take the range of n to be $(-\infty, +\infty)$ keeping in mind that the approximation only makes sense when N is sufficiently large, if $0 < \eta < 1$, and η is not too close to either 0 or 1.

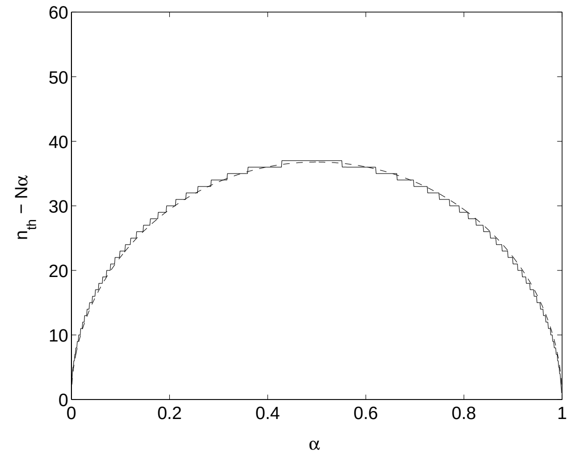

FIG. 10: Graph of $n_{\rm th} - N\alpha$ versus the false alarm probability $\alpha = e^{-\rho_{\rm th}}$ for $\alpha_H = 0.01$ and $N = 2000$. The dashed line shows the analytic approximation given by equation (5.22).

With the approximations given above, we can rewrite the equation $\alpha_H = \alpha_H^\star$ as

$$\int_{n_{\rm th}}^{\infty} p(n|\rho_{\rm th}, 0)\, dn = \alpha_H^\star \,. \tag{5.20}$$

This equation can be rewritten in terms of the complementary error function:

$$\mathrm{erfc}\left(\frac{n_{\rm th} - N\alpha}{\sqrt{2}\sigma}\right) = 2\alpha_H^\star \,. \tag{5.21}$$

Thus, given $\alpha_H^\star$, the solution for $n_{\rm th}$ is:

$$\begin{aligned} n_{\rm th}(\rho_{\rm th}, N, \alpha_H^\star) &= N\alpha + \sqrt{2}\sigma\ \mathrm{erfc}^{-1}(2\alpha_H^\star) \\ &= N\alpha + \sqrt{2N\alpha(1-\alpha)}\ \mathrm{erfc}^{-1}(2\alpha_H^\star) \,. \end{aligned} \tag{5.22}$$

As shown in figure 10, this is a very good approximation to the actual value of $n_{\rm th}$ obtained from the binomial distribution.

The expression for β_H is similarly rewritten as

$$\begin{aligned} \beta_H &= \int_{-\infty}^{n_{\rm th}} p(n|\rho_{\rm th}, \lambda)\, dn \\ &= \frac{1}{2}\mathrm{erfc}\left(\frac{N\eta - n_{\rm th}}{\sqrt{2N\eta(1-\eta)}}\right) \,. \end{aligned} \tag{5.23}$$

As figure 11 shows, this too is a very good approximation to β_H obtained using the binomial distribution.

In step (iv), we find $\rho_{\rm th}^\star$ such that

$$\left.\frac{\partial \beta_H}{\partial \rho_{\rm th}}\right|_{\rho_{\rm th}^\star} = 0 \tag{5.24}$$

which, using equation (5.23), leads to

$$\left.\frac{\partial}{\partial \rho_{\rm th}}\left(\frac{N\eta - n_{\rm th}}{\sqrt{\eta(1-\eta)}}\right)\right|_{\rho_{\rm th}^\star} = 0 \,. \tag{5.25}$$

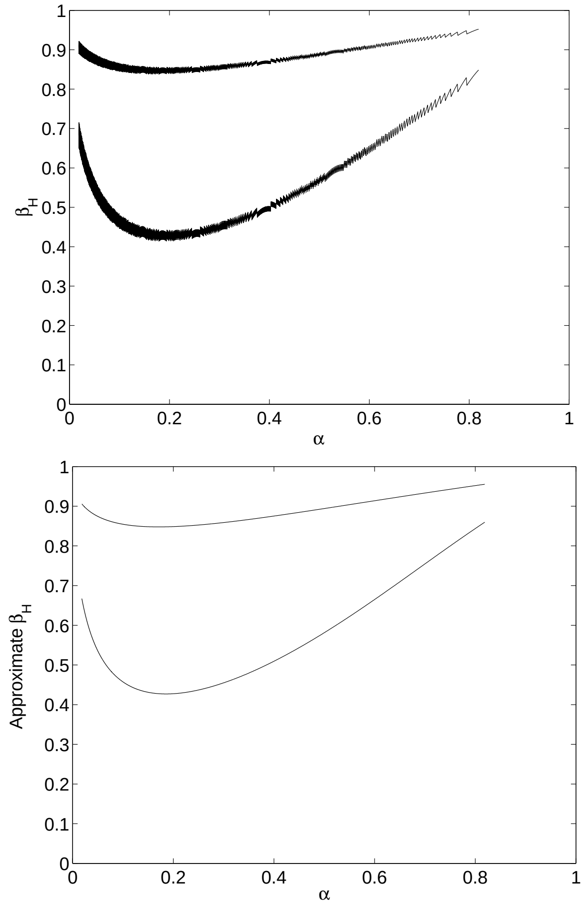

FIG. 11: The first figure shows the Hough false dismissal rate as a function of the false alarm rate $\alpha = e^{-\rho_{\rm th}}$ for a non-centrality parameter $\lambda = 0.10$ (upper curve) and $\lambda = 0.20$ (lower curve). Both curves correspond to $\alpha_H = 0.01$ and $N = 1000$. The minimum values of β_H for the two curves are approximately 0.84 and 0.41 respectively. Both minima occur at $\alpha = 0.20$ approximately. This corresponds to a threshold of $\rho_{\rm th} = 1.6$ on the normalized power statistic. The bottom figure shows the approximation to β_H using equations (5.23) and (5.22) with the same parameters as in the first figure.

Let us now specialize to the case of small signals where $\eta \approx \alpha(1+\rho_{\rm th}\lambda/2)$. In this case, the solution to the above equation becomes independent of λ and is given by

$$\frac{\partial}{\partial\rho_{\rm th}}\left(\sqrt{\frac{N\rho_{\rm th}^2}{e^{\rho_{\rm th}}-1}} + \frac{{\rm erfc}^{-1}(2\alpha_H^\star)\rho_{\rm th}}{2}\frac{1-2e^{-\rho_{\rm th}}}{1-e^{-\rho_{\rm th}}}\right)\Bigg|_{\rho_{\rm th}^\star} = 0 \tag{5.26}$$

Note that the second term within brackets in the above equation is independent of N whence the first term dominates for large N:

$$\frac{\partial}{\partial\rho_{\rm th}}\left(\sqrt{\frac{N\rho_{\rm th}^2}{e^{\rho_{\rm th}}-1}}\right)\Bigg|_{\rho_{\rm th}^\star} = 0\,. \tag{5.27}$$

The solution to this equation is $\rho_{\rm th}^\star \approx 1.6$ and $\alpha^\star = e^{-\rho_{\rm th}^\star} \approx 0.20$. Notice that this equation is equivalent

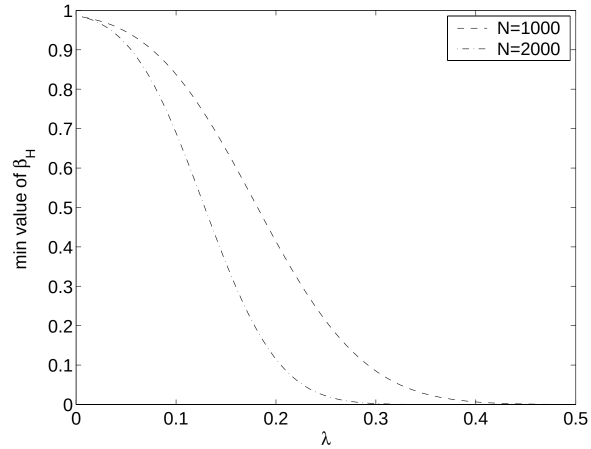

FIG. 12: Minimum value of β_H as a function of the non-centrality parameter λ for $\alpha_H = 0.01$ and for $N = 1000$ and 2000. As expected, a larger value of N typically leads to a smaller value of β_H.

to equation (5.17) and furthermore, the functions being extremized are rather flat near the extremum. Thus, the threshold could be chosen differently without significantly impacting the sensitivity. In particular, the threshold can be increased so that fewer frequency bins are selected. Depending on the details of the implementation, this could lead to a lower computational cost; in the framework of a hierarchical search, this will improve the overall sensitivity.

Finally, with the optimal threshold $\rho_{\rm th}^\star$ at hand, the optimal threshold $n_{\rm th}^\star$ on the number count is obtained by substituting $\rho_{\rm th} = \rho_{\rm th}^\star$ in equation (5.22):

$$\begin{aligned} n_{\rm th}^\star &= n_{\rm th}(\rho_{\rm th}^\star, N, \alpha_H^\star) \\ &= N\alpha^\star + \sqrt{2N\alpha^\star(1-\alpha^\star)}\,{\rm erfc}^{-1}(2\alpha_H^\star)\,. \end{aligned} \tag{5.28}$$

In performing the analysis, this is an important equation because it tells us the number count threshold that must be set in order to a have a given number of follow-up candidates.

C. Sensitivity

In this section, we estimate the sensitivity of the Hough search, i.e. we answer the following question: for given values $\alpha_H^\star$ and $\beta_H^\star$ of the false alarm α_H and false dismissal β_H respectively, what is the smallest value of the gravitational wave amplitude h_0 (see equation (2.9)) that would cross the thresholds $\rho_{\rm th}$ and $n_{\rm th}$? Equivalently, for a given false alarm rate $\alpha_H^\star$, what is the smallest h_0 which will give a false dismissal rate of at least $\beta_H^\star$? For concreteness, we use the signal model of equation (2.7) with and we will present our final result for the values $\alpha_H = \alpha_H^\star = 0.01$ and $\beta_H = \beta_H^\star = 0.10$. The value of 0.01 is meant mainly for illustration purposes and does not change the results qualitatively. Furthermore, the

S1 analysis [5] uses a false alarm of 0.01 and this choice of $\alpha_H^\star$ enables an easier comparison of results. As far as possible, we explicitly retain the factors of $\alpha_H^\star$ in our equations substituting numerical values only when necessary.

We must first solve the equation

$$\beta_H(n_{\rm th}^\star, \rho_{\rm th}^\star, \lambda, N) = \beta_H^\star \tag{5.29}$$

and obtain λ as a function of N; this will yield the desired value of h_0.

In order to simplify the discussion, we shall again approximate the binomial distribution by a normal distribution whose mean $\bar{n}$, and variance σ, are respectively given by equation (5.11). The false dismissal rate is then given by

$$\begin{aligned} \beta_H &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_0^{n_{\rm th}^\star} e^{-(n-\bar{n})^2/2\sigma^2}\, dn \\ &\approx \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{n_{\rm th}^\star} e^{-(n-\bar{n})^2/2\sigma^2}\, dn \\ &= \frac{1}{2}\mathrm{erfc}\left(\frac{\bar{n} - n_{\rm th}^\star}{\sqrt{2}\sigma}\right) \end{aligned} \tag{5.30}$$

where $n_{\rm th}^\star$ is as given in equation (5.22). More explicitly, this can be written as

$$\beta_H = \frac{1}{2}\mathrm{erfc}\left(\frac{N\eta^\star - N\alpha^\star - \sqrt{2N\alpha^\star(1-\alpha^\star)}\,\mathrm{erfc}^{-1}(2\alpha_H^\star)}{\sqrt{2N\eta^\star(1-\eta^\star)}}\right) \tag{5.31}$$

where $\eta^\star = \eta(\rho_{\rm th}^\star|\lambda)$.

Since we are interested in the case of small signals, let us approximate η by only keeping terms of the order of λ in equation (5.9). Ignoring terms of $\mathcal{O}(\lambda^2)$, equation (5.31) leads to the approximation

$$\beta_H = \frac{1}{2}\mathrm{erfc}\left(-\mathrm{erfc}^{-1}(2\alpha_H^\star) + \frac{1}{2}\Theta\alpha^\star\rho_{\rm th}^\star\lambda\right). \tag{5.32}$$

where

$$\Theta = \sqrt{\frac{N}{2\alpha^\star(1-\alpha^\star)}} + \left(\frac{1-2\alpha^\star}{1-\alpha^\star}\right)\frac{\mathrm{erfc}^{-1}(2\alpha_H^\star)}{2\alpha^\star} \tag{5.33}$$

Let us summarize our approximation scheme for β_H. The first approximation is to take the number count distribution to be binomial. The second approximation is in equation (5.31) which replaces the binomial by a Gaussian distribution with the appropriate mean and variance. The final approximation is in equation (5.32) where we have taken λ to be small and used a Taylor series in powers of λ retaining only the linear term. To get a feeling for the validity of these approximations, figure 13 shows graphs of β_H as a function of λ for different values of N. As the graphs show, we can trust the approximations when $N \sim 10^3$. For smaller values of N, while the Gaussian approximation is still reasonable, the linear approximation greatly under-estimates β_H for a given value of λ, i.e. it make the Hough search appear much more sensitive than it actually is.

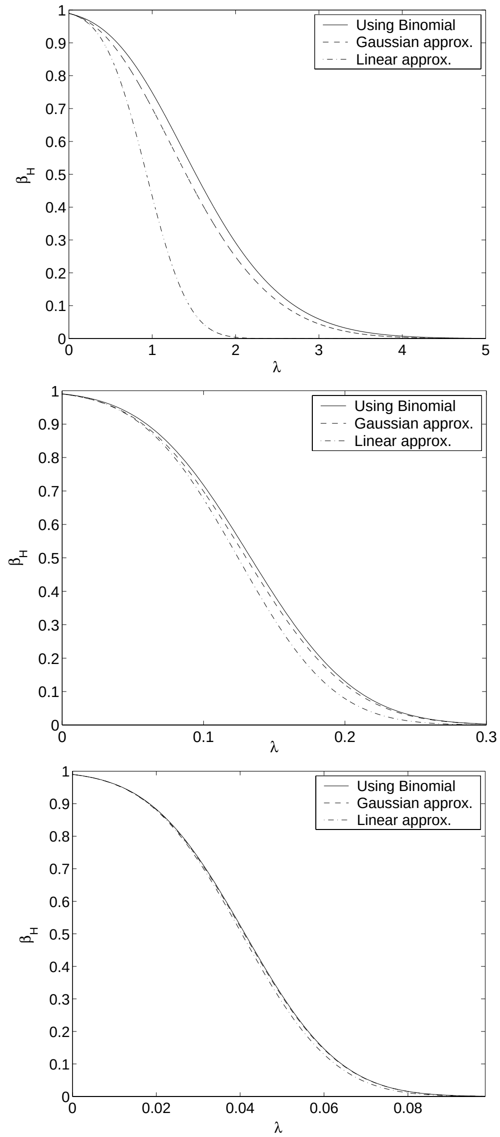

FIG. 13: Graphs of β_H as a function of λ for different values of N and for the three different approximations used. In the first panel $N = 20$, the second panel has $N = 2000$ and in the third panel, $N = 20,000$. All graphs are plotted assuming the optimal values for $\rho_{\rm th}$ and $n_{\rm th}$. The linear approximation is clearly unacceptable for $N \sim 10^1$ but becomes reasonable when $N \sim 10^2$ or 10^3 and is excellent for $N \sim 10^4$. The Gaussian approximation is clearly much better and is good even for $N = 20$. Finally, note that the approximations always underestimate the value of β_H.

Working with the linear approximation of equation (5.32), assuming N to be very large and $\mathrm{erfc}^{-1}(2\alpha_H^\star) \ll N$, set $\beta_H = \beta_H^\star$ and solve for λ:

$$\lambda \approx \frac{\mathcal{S}}{\rho_{\rm th}^\star}\sqrt{\frac{8(1-\alpha^\star)}{N\alpha^\star}} \approx \frac{9.02}{\sqrt{N}} \tag{5.34}$$

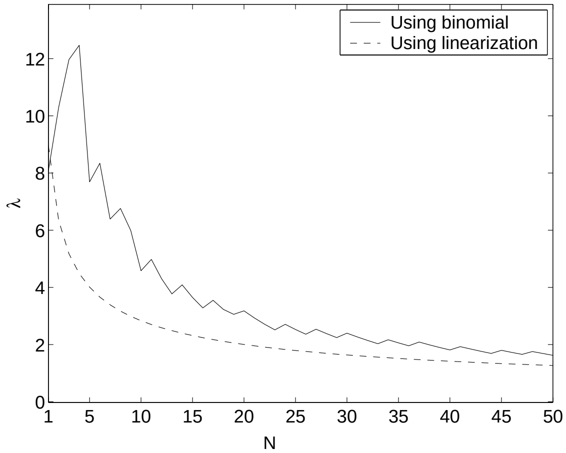

FIG. 14: Graph of the smallest detectable λ with the optimal thresholds. The dashed curve uses the linear approximation of equation (5.32) while the solid curve uses the binomial distribution. See text for additional discussion.

where

$$\mathcal{S} := \mathrm{erfc}^{-1}(2\alpha_H^\star) + \mathrm{erfc}^{-1}(2\beta_H^\star) \qquad (5.35)$$

and to obtain numerical values, we have chosen $\alpha_H^\star = 0.01$ and $\beta_H^\star = 0.10$. Using the properties of the complemetary error function, it is easy to show that $\mathcal{S} = 0$ implies that the statistical significance $s := 1 - \alpha_H^\star - \beta_H^\star$ also vanishes. Therefore, the quantity $\mathcal{S}$ can be taken to be a measure of the statistical significance of the search. The value of λ obtained in equation (5.34) gives us the strength of the smallest signal that can be detected by the Hough search with a false alarm rate of 1% and a false dismissal rate of 10%.

A graph of λ as a function of N for small values of N is shown in figure (14); this figure shows the results using both the linear approximation and the more accurate binomial distribution. The small N limit requires a brief explanation. For small N, the discrete nature of n becomes important. In particular, the false alarm α_H defined in equation (5.13) can have only a discrete number of values, the smallest of which is α^N at $n_{\rm th} = N$. Thus for $N = 1$, it is not possible to reach the desired 1% false alarm rate and the best we can do is $\alpha_H = 0.2$. To find the value of λ which yields $\beta_H = 0.1$, note that here $\beta_H = 1 - \eta$, and $\eta = 0.9$ implies $\lambda \approx 8.08$; this then is the sensitivity of the search for $N = 1$. It corresponds to a false alarm rate of 20% and a false dismissal rate of 1%. Similar calculations show that the sensitivity becomes worse as N is increased from 1 to 4 as the corresponding false alarm rates become better. It is only at $N = 5$ that we can choose $n_{\rm th} < N$ from which point onwards the sensitivity begins to improve. This explains the small N behavior of figure 14. Similarly, the other discrete jumps in figure 14 are due to the discrete nature of α_H and requirement of keeping it below the 1% level.

To recast the expression for λ directly in terms of the signal amplitude, start with equation (5.3); λ depends on the various pulsar parameters. The relevant quantity for the purposes of this section is the *average* of λ over these parameters. It is quite straightforward to estimate this average. First, recall the expression for $h(t)$:

$$h(t) = F_+(t) A_+ \cos\Phi(t) + F_\times A_\times \sin\Phi(t)\,. \qquad (5.36)$$

Since $T_{\rm coh}$ is much lesser than a day and much larger than the period of the pulsar, we can take $F_{+\times}$ and the signal frequency $f(t)$ to be roughly constant. Thus, we get the expression for the fourier component of the signal:

$$\tilde{h}(f_k) \approx \frac{T_{\rm coh}}{2}(F_+A_+ + F_\times A_\times)\frac{\sin[\pi(f-f_k)T_{\rm coh}]}{\pi(f-f_k)T_{\rm coh}} \qquad (5.37)$$

where f is the instantaneous frequency of the signal and f_k is the central Fourier frequency of the frequency bin containing f; f is allowed to lie in the range $(f_k - \delta f/2, f_k + \delta f/2)$. Now take the square of $\tilde{h}_k$ and average over time to get the average non-centrality parameter for all the SFTs and note that the time averages F_+^2 and $F_\times^2$ are both 1/5 and the time average of $F_+F_\times$ vanishes. Thus:

$$\lambda \approx \frac{T_{\rm coh}}{10 S_n}(A_+^2 + A_\times^2)\left(\frac{\sin[\pi(f-f_k)T_{\rm coh}]}{\pi(f-f_k)T_{\rm coh}}\right)^2 . \qquad (5.38)$$

Take the amplitudes to be of the form given in equations (2.7) and average over $\cos\iota \in (-1,1)$ and over the values of the signal frequency $f \in (f_k - \delta f/2, f_k + \delta f/2)$:

$$\begin{aligned}\langle\lambda\rangle_{\iota,\psi,f,\alpha,\delta} &\approx \frac{4}{25}\frac{h_0^2 T_{\rm coh}}{S_n}\left(\int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\sin^2(\pi x)}{(\pi x)^2}\,dx\right) \\ &\approx 0.7737 \times \frac{4}{25}\frac{h_0^2 T_{\rm coh}}{S_n}\,. \end{aligned} \qquad (5.39)$$

Therefore, we get the following value for the smallest signal that can be detected by the Hough search:

$$h_0 = \frac{8.54}{N^{1/4}}\sqrt{\frac{S_n}{T_{\rm coh}}} = 8.54 N^{1/4}\sqrt{\frac{S_n}{T_{\rm obs}}}\,. \qquad (5.40)$$

This is the result we were looking for. This tells us that if we wish to detect a signal with a false alarm rate of 1% and a false dismissal rate of 10%, the weakest signal that will cross the optimal thresholds is the h_0 given above. The important feature to note is that h_0 is proportional to $N^{1/4}/\sqrt{T_{\rm obs}}$ while for a coherent search over the whole observation time, the sensitivity is proportional to $1/\sqrt{T_{\rm obs}}$. In particular, for the same values of the false alarm and false dismissal rates as above, the sensitivity of a full coherent search directed at around a single point in parameter space is given by (see [5]):

$$h_0 = 11.4\sqrt{\frac{S_n}{T_{\rm obs}}}\,. \qquad (5.41)$$

This illustrates the loss in sensitivity introduced by combining the different SFTs incoherently but, of course, this is compensated by the lesser computational requirements for the incoherent method. Furthermore, for say $N = 2000$, the sensitivity of the Hough search is only about a factor of 4.5 worse than a full directed coherent search.

This result helps one to make trade-offs of coherent against hierarchical searches. For example, if one is searching for a population of objects that is uniformly distributed in a plane, such as a population of young pulsars in the Galaxy, then a coherent search of any region of parameter space would go 4.5 times deeper than the incoherent method with $N = 2000$. The volume of space surveyed would be $4.5^2 = 20$ times larger. However, if the incoherent method's speed of execution allowed it to survey more than 20 times as much parameter space (including sky area and spindown range) then one would choose the incoherent method. This is indeed the case for pulsar searches.

Finally, equation (5.40) also allows us to estimate the astrophysical range of the search. Combining (2.9) and (5.40), we get:

$$
\begin{aligned}
d &= \frac{16\pi^2 G N^{1/4} I_{zz}\epsilon f_r^2}{8.54c^4}\sqrt{\frac{T_{\rm coh}}{S_n(f_r)}} \qquad (5.42)\\
&= 5.8\,\mathrm{kpc}\times\left(\frac{N}{17000}\right)^{\frac{1}{4}}\left(\frac{I_{zz}}{10^{38}\mathrm{kg\text{-}m}^2}\right)\left(\frac{f_r}{500\mathrm{Hz}}\right)^2\\
&\quad\times\left(\frac{\epsilon}{10^{-6}}\right)\left(\frac{T_{\rm coh}}{1800\mathrm{s}}\right)^{\frac{1}{2}}\left(\frac{10^{-46}\mathrm{Hz}^{-1}}{S_n}\right)^{\frac{1}{2}}.
\end{aligned}
$$

Here we have taken a total observation time of approximately 1yr so that, with $T_{\rm coh} = 1800s$, we would have $N \approx 17000$, and we have taken the detector sensitivity to be $10^{-23}\mathrm{Hz}^{-1/2}$ at a frequency of 1kHz which is appropriate for the proposed advanced LIGO detector[23]. The biggest uncertainty in this equation is in the value of the pulsar ellipticity ϵ.

VI. HOUGH TRANSFORM WITH DEMODULATED DATA

As equation (4.10) shows, using the Hough transform with SFTs as input necessarily limits the coherent time-baseline $T_{\rm coh}$, and therefore also the sensitivity of the search. To get around this limitation, we need to demodulate each coherent data segment to remove the frequency drifts caused by the Doppler modulation and the spindown; the only limitation on $T_{\rm coh}$ is then due to the available computational resources. The demodulation procedure we use is based on the $\mathcal{F}$-statistic introduced in [6] and the search pipeline is shown in figure 15. Section VI A provides a brief description of the $\mathcal{F}$-statistic, the master equation is derived in section VI B, section VI C provides the implementation details and section VI D describes the statistics.

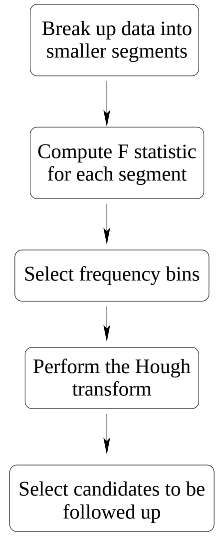

FIG. 15: Data analysis pipeline for a Hough transform search using the $\mathcal{F}$ statistic.

A. The $\mathcal{F}$ statistic

Let $x(t)$ be the calibrated detector output and let $h(t)$ the waveform that we are searching for. In order to extract the signal $h(t)$ from the noise, the optimal search statistic is the likelihood function Λ defined by

$$\ln\Lambda = (x|h) - \frac{1}{2}(h|h) \qquad (6.1)$$

where the inner product $(\cdot|\cdot)$ is defined as

$$(x|y) := 2\int_0^\infty \frac{\tilde{x}(f)\tilde{y}^\star(f) + \tilde{x}^\star(f)\tilde{y}(f)}{S_n(f)}\,\mathrm{d}f\,. \qquad (6.2)$$

Here, as before, $\tilde{x}(f)$ is the Fourier transform of $x(t)$ and $S_n(f)$ is the one-sided power spectral density. The expected waveform $h(t)$ is given by equations (2.6), (2.2) and (2.3). The quantity $\ln\Lambda$ is essentially the matched filter and is precisely what we should use in order to best detect the waveform $h(t)$. However, apart from the parameters $\vec{\xi} = (f_{(0)}, \{f_{(n)}\}, \mathbf{n})$, $\ln\Lambda$ also depends upon the other parameters such as the orientation of the pulsar, the polarization angle of the wave etc. The $\mathcal{F}$-statistic eliminates these additional variables and enables us to search over only the shape parameters $\vec{\xi}$.

Following the notation of [6], the dependence of the antenna patterns $F_{+,\times}$ on the polarization angle ψ are given by

$$
\begin{aligned}
F_+(t) &= \sin\zeta[a(t)\cos 2\psi + b(t)\sin 2\psi] \qquad (6.3)\\
F_\times(t) &= \sin\zeta[b(t)\cos 2\psi - a(t)\sin 2\psi] \qquad (6.4)
\end{aligned}
$$

where the functions $a(t)$ and $b(t)$ are independent of ψ and ζ is the angle between the arms of the detector. If we write the phase of the gravitational wave as

$$\Phi(t) = \phi_0 + \phi(t)\,, \qquad (6.5)$$

then we can always decompose the total waveform $h(t)$ in terms of four quadratures as

$$h(t) = \sum_{i=1}^{4} A_i h_i(t) \tag{6.6}$$

where the four amplitudes A_i are time independent and the h_i are as follows:

$$\begin{aligned} h_1(t) &= a(t)\cos\phi(t)\,, \qquad h_2 = b(t)\cos\phi(t)\,, \\ h_3(t) &= a(t)\sin\phi(t)\,, \qquad h_4 = b(t)\sin\phi(t)\,. \end{aligned} \tag{6.7}$$

What this decomposition achieves is a separation of the shape parameters $\vec{\xi}$ from the other pulsar parameters. The *only* unknown parameters in the quadratures h_i are the shape parameters $\vec{\xi}$ while the amplitudes A_i are independent of $\vec{\xi}$. The log likelihood function depends quadratically on the four amplitudes and we can analytically find the maximum likelihood (ML) estimators $\hat{A}_i$ of the amplitudes A_i by solving the set of four coupled linear equations

$$\left.\frac{\partial \ln \Lambda}{\partial A_i}\right|_{A_i=\hat{A}_i} = 0\,, \qquad i = 1,\ldots,4\,. \tag{6.8}$$

The $\mathcal{F}$ statistic is then defined as the log likelihood ratio with the values of the amplitudes A_i replaced by their ML estimators:

$$\mathcal{F} := \ln \Lambda|_{A_i=\hat{A}_i}\,. \tag{6.9}$$

The only unknown parameters in the optimal search statistic $\mathcal{F}$ are the shape parameters $\vec{\xi}$.

B. The master equation

Equation (2.4) describes the expected time-frequency pattern when the search statistic is the Fourier transform; in other words, if the detector output $x(t)$ contains a true signal with instantaneous frequency $\hat{f}(t)$, then equation (2.4) tells us the value of the observed frequency $f(t)$ which would maximise $\tilde{x}(f)$ in the absence of noise. If we now use the $\mathcal{F}$-statistic instead of the Fourier transform, the expected time-frequency pattern is, as described below, different.

Before proceeding further, it is useful to distinguish the instantaneous frequency $f_{(0)}$ from the other shape parameters which we denote by $\vec{\lambda}$: $\vec{\xi} = (f_{(0)}, \vec{\lambda}) = (f_{(0)}, \{f_{(n)}\}, \mathbf{n})$. Let us the assume that the $\mathcal{F}$ statistic has been computed using the parameters $\vec{\lambda}_d$ but that the detector output consists of a signal with parameters $(f_{(0)}, \vec{\lambda})$; let us denote the mismatch in the parameters by $\Delta\vec{\lambda} := \vec{\lambda} - \vec{\lambda}_d$ and $\Delta f := f - f_{(0)}$.

Since the $\mathcal{F}$-statistic is maximized when the source parameters are equal to the demodulation parameters, it is clear that if $\Delta\vec{\lambda} = 0$, then the expected time-frequency pattern is just a constant frequency $f = f_{(0)}$. More generally, due to the correlations in parameter space, the mismatch $\Delta\vec{\lambda}$ may produce a residual shift in the frequency Δf. This frequency shift is determined by

$$\left.\frac{\partial \mathcal{F}(f, \vec{\lambda}_d; f_{(0)}, \vec{\lambda})}{\partial f}\right|_{\vec{\lambda}_d, f_{(0)}, \vec{\lambda}} = 0\,. \tag{6.10}$$

Expand $\mathcal{F}$ in powers of $\Delta\vec{\lambda}$ and Δf around the point $(f_{(0)}, \vec{\lambda})$ upto second order (repeated indices are summed over):

$$\begin{aligned} \mathcal{F}(f, \vec{\lambda}_d; f_{(0)}, \vec{\lambda}) &= \mathcal{F}(f_{(0)}, \vec{\lambda}) + A_{00}\,(\Delta f)^2 \\ &+ A_{0i}\Delta\lambda_i \Delta f + A_{ij}\Delta\lambda_j\Delta\lambda_j\,. \end{aligned} \tag{6.11}$$

The linear terms do not appear because $\mathcal{F}$ is maximized when $\Delta f = 0$ and $\Delta\vec{\lambda} = 0$. With this approximation, equation (6.10) leads to the master equation

$$\Delta f = \frac{A_{0i}}{2A_{00}}\Delta\lambda_i \tag{6.12}$$

In other words, the frequency value that maximizes the $\mathcal{F}$ statistic for a given $\Delta\vec{\lambda}$, does not correspond to the intrinsic source frequency $f_{(0)}$ but is instead given by a linear combination of the $\Delta\lambda_i$.

Let us rewrite equation (6.10) more explicitly. As shown in [6], $\mathcal{F}$ can be written in terms of the amplitude modulation functions $a(t)$ and $b(t)$ as

$$\mathcal{F} = \frac{4}{T_{\rm coh} S_n(f_{(0)})} \frac{B|F_a|^2 + A|F_b|^2 - 2C\mathcal{R}(F_a F_b^\star)}{D} \tag{6.13}$$

where A, B, C, and D are constants and

$$F_a = \int_{-T_{\rm coh}/2}^{T_{\rm coh}/2} x(t)a(t)e^{-i\Phi(t;f,\vec{\lambda}_d)}\,dt\,, \tag{6.14}$$

$$F_b = \int_{-T_{\rm coh}/2}^{T_{\rm coh}/2} x(t)b(t)e^{-i\Phi(t;f,\vec{\lambda}_d)}\,dt\,. \tag{6.15}$$

Since we are interested in calculating the frequency drift and not the amplitude, the variation in the phase is more important than the amplitude modulation; the factors of $a(t)$ and $b(t)$ can be ignored in the above equation. Thus, maximizing $\mathcal{F}$ is equivalent to maximizing $|\tilde{X}(f)|^2$ where $\tilde{X}(f)$ is the *demodulated Fourier transform* (DeFT) defined as

$$\tilde{X}(f) = \int x(t)e^{-i\Phi(t;f,\vec{\lambda}_d)}\,dt\,. \tag{6.16}$$

With this approximation, the master equation is obtained by solving

$$\left.\frac{\partial |\tilde{X}(f, \vec{\lambda}_d; f_{(0)}, \vec{\lambda})|^2}{\partial f}\right|_{\vec{\lambda}_d, f_{(0)}, \vec{\lambda}} = 0\,. \tag{6.17}$$

The details of the calculation are described in appendix A; the result is:

$$f(t) - F_0(t) = \vec{\zeta}(t) \cdot (\mathbf{n} - \mathbf{n}_d) \tag{6.18}$$

where

$$F_0(t) = f_{(0)} + \sum_{k=1}^{\infty} \frac{\Delta f_{(k)}}{k!} (\Delta t)^k \ , \tag{6.19}$$

and

$$\begin{aligned} \vec{\zeta}(t) \ &= \ \left(F_0(t) + \sum_{k=1}^{k=\infty} \frac{f_{d(k)}}{k!} (\Delta t)^k \right) \frac{\mathbf{v}(t)}{c} \\ &+ \ \left(\sum_{k=1}^{k=\infty} \frac{f_{d(k)}}{(k-1)!} (\Delta t)^{k-1} \right) \frac{\mathbf{r}(t) - \mathbf{r}(t_0)}{c} \ . \end{aligned} \tag{6.20}$$

The $\Delta f_{(k)}$'s are the residual spindown parameters: $\Delta f_{(k)} = f_{(k)} - f_{d(k)}$. As expected, if $\Delta\vec{\lambda} = 0$ so that $\mathbf{n} = \mathbf{n}_d$ and $f_{(k)} = f_{d(k)}$, then $f(t) = f_{(0)}$. Furthermore, it is clear that this master equation is *qualitatively* similar to equation (2.4) except for a constant frequency offset $\vec{\zeta} \cdot \mathbf{n}_d$. Thus, many of the methods obtained for the non-demodulated case will still be valid.

C. Implementation details

As mentioned above, the master equations (2.4) and (6.18) are qualitatively similar except for a constant frequency offset. Thus, many of the earlier results are still valid with some minor modifications which we now explain.

Resolution in parameter space: The formula for the resolution in f_k space is the same as given in equation (4.16). However, since we can make $T_{\rm coh}$ much larger than before, the resolution can be made much more finer. Thus, for the first spindown parameter, instead of equation (4.17), we would have

$$\delta f_{(1)} = (1.3 \times 10^{-13}\mathrm{Hz/s}) \times \frac{1000}{N} \left(\frac{1\mathrm{day}}{T_{\rm coh}} \right)^2 . \tag{6.21}$$

Furthermore, the restriction due to the length of $T_{\rm coh}$ (see equation (4.21)) is no longer an issue.

As for the sky-positions, using the approximation given in equation (6.24), the estimate of the resolution in the sky proceeds in the same way as the derivation of equation (4.12). The results of equations (4.15) and (4.16) are still valid, the only change being that $T_{\rm coh}$ is now of the order of a day. Therefore, we rewrite equations (4.15) and (4.16) as:

$$\begin{aligned} (\delta\phi)_{\rm min} &= \frac{c}{v}\frac{\delta f}{\hat{f}} = \frac{c}{v\hat{f}T_{\rm coh}} \\ &= 1.0 \times 10^{-3}\,\mathrm{rad} \times \left(\frac{1\mathrm{day}}{T_{\rm coh}} \right) \left(\frac{500\,\mathrm{Hz}}{\hat{f}} \right) \left(\frac{10^{-4}}{v/c} \right) \end{aligned} \tag{6.22}$$

and

$$\begin{aligned} (\delta\phi)_{\rm max} &= \sqrt{\frac{\delta f}{\hat{f}}\frac{c}{v}} = \sqrt{\frac{c}{v\hat{f}T_{\rm coh}}} \\ &= 1.5 \times 10^{-2}\,\mathrm{rad} \times \left(\frac{1\mathrm{day}}{T_{\rm coh}} \right)^{\frac{1}{2}} \left(\frac{500\mathrm{Hz}}{\hat{f}} \right)^{\frac{1}{2}} \left(\frac{10^{-4}}{v/c} \right)^{\frac{1}{2}} . \end{aligned} \tag{6.23}$$

The sky resolution is therefore about 5 times better than in the non-demodulated case.

Sky-patch size: Unlike in the non-demodulated case, since we are removing the frequency modulation of the signal beforehand, there is now, except for computational constraints, no restriction at all on the coherent integration time $T_{\rm coh}$. Typically, $T_{\rm coh}$ will be taken to be of the order of a day. However, the price we pay for this is that the demodulation is not valid for arbitrarily large patches. The *patch size* is determined by the largest fractional loss of sensitivity (e.g., the $\mathcal{F}$ value) we are willing to tolerate from a putative source with a certain mismatch parameters $\Delta\vec{\xi}$ and this restricts both the sky location and the spindown parameters.

If we have demodulated for a direction $\mathbf{n}_d$ in the sky, how different can $\mathbf{n}$ be from $\mathbf{n}_d$ so that the loss in the signal power does not become unacceptably large? In order to answer this question, we would have to analyze the parameter space metric defined in terms of the mismatch [22]. The analysis of the metric will be presented elsewhere, but in this paper we shall just use a conservative estimate for the size of the sky-patch.

To estimate the size of the sky-patch, first note that the quantity ξ is, to a very good approximation, given by

$$\zeta(t) \approx \hat{f}\frac{\mathbf{v}(t)}{c} \tag{6.24}$$

where, as before, $\mathbf{v}$ is the velocity of the detector in the SSB frame. The velocity $\mathbf{v}(t)$ is the sum of two components, the velocity $\mathbf{v}_y$ due to the yearly motion around the sun and the velocity $\mathbf{v}_d$ due to the rotation of Earth around its axis: $\mathbf{v} = \mathbf{v}_y + \mathbf{v}_d$. For reference, note that for the GEO detector, the magnitude v_y is about 10^2 times larger than v_d. The estimate of the sky-patch size proceeds roughly like the estimate of the pixel size in equation (6.23) except for one difference. If we take the coherent integration time $T_{\rm coh}$ to be roughly of the order of less than a day, say a third of a day, then the relevant velocity is $\mathbf{v}_d$. Thus, the sky-patch size h is roughly given by

$$h \approx \frac{c}{v_d}\frac{\delta f}{\hat{f}} = \frac{c}{v_d\hat{f}T_{\rm coh}} \ . \tag{6.25}$$

Since v_d is roughly 100 times smaller than v, $h \approx 100(\delta\phi)_{\rm min}$. Thus, a typical sky-patch consists of about 100 pixels on a side. It should be emphasized that this is only an educated guess and is not likely to be valid for larger $T_{\rm coh}$. A more complete analysis would require the calculation of the metric on parameter space.

Validity of the look up tables: Again using the approximation given in equation (6.24), the number of frequency bins for which the LUT is valid can be estimated in a similar way as in the non-demodulated case. The master equation is

$$\Delta f := f - F_0 = \hat{f}\frac{\mathbf{v}}{c}\cdot(\mathbf{n}-\mathbf{n}_d) \qquad (6.26)$$

Rewrite the equation as

$$\frac{1}{\hat{f}} = \frac{1}{\Delta f}\frac{v}{c}\cos\phi - \frac{1}{\Delta f}\frac{\mathbf{v}\cdot\mathbf{n}_d}{c}\,. \qquad (6.27)$$

Keeping Δf fixed and differentiating w.r.t. $\hat{f}$ leads to

$$\frac{d\hat{f}}{\hat{f}^2} = \frac{1}{\Delta f}\frac{v}{c}\sin\phi\, d\phi \qquad (6.28)$$

As before, define κ and r by $d\hat{f} = \kappa\delta f$ and $d\phi = r\,(\delta\phi)_{\rm min}$.

Substituting these definitions in the above equation yields

$$\kappa = \frac{r f_0}{\Delta f}\sin\phi \qquad (6.29)$$

There are now two cases to look at, namely when ϕ is close to $\pi/2$ or when it is close to 0 (or π). First the easy case when $\phi\sim\pi/2$. Here the width of the annuli is roughly the same as the pixel size: $\delta\phi\sim(\delta\phi)_{\rm min}$. Thus, if h is the length of a side of the sky-patch (assumed to be square) then the number of annuli in the sky patch is $h/\delta\phi$ which means $\Delta f = \delta f\cdot(h/\delta\phi)$. Substituting this in equation (6.29) and also setting $\phi=\pi/2$ finally leads to the result

$$\kappa = \kappa_0 := \frac{r}{hv/c}\,. \qquad (6.30)$$

Now turn to the large annulus case. The annulus size is given by $\delta\phi\sim((\delta\phi)_{\rm max})$ and again $\Delta f = \delta f\cdot(h/\delta\phi)$. As for the numerator of equation (6.30), take the smallest value of $\sin\phi$, i.e. when ϕ is no bigger than a pixel so that $\sin\phi\sim(\delta\phi)_{\rm min}$. Substituting these estimates leads to

$$\kappa = \left(\frac{r}{hv/c}\right)\sqrt{\frac{\delta f}{f_0}\frac{c}{v}} = \kappa_0\,(\delta\phi)_{\rm max}\,. \qquad (6.31)$$

From equations (6.24) and (6.30) we see that typically, κ for the thick annulus case is about 100 times smaller than for the thin annulus case.

D. Statistics

This section discussion describes the statistics of the Hough map and the $\mathcal{F}$-statistic, the optimal thresholds and the sensitivity. The discussion closely parallels that of section V; here we simply point out some of the differences that arise when the $\mathcal{F}$-statistic is considered instead of the normalized power.

Just as the distribution of the normalized ρ_k power in section V A turned out to be related to the χ^2 distribution with 2 degrees of freedom, one might intuitively expect that the distribution of $\mathcal{F}$ should also be related to a χ^2 distribution. However, since $\mathcal{F}$ is constructed from the four filters given in equation (6.7), it turns out that the distribution of $2\mathcal{F}$ is a non-central χ^2 distribution with *four* degrees of freedom. As before, we shall denote the non-centrality parameter by λ, and it turns out to be

$$\lambda = (h|h) \qquad (6.32)$$

where the inner product $(\cdot|\cdot)$ has been defined in equation (6.2). Note that while we use the same symbol for the non-centrality parameter as in the non-demodulated case, this definition is different from that of equation (5.3).

Thus, the distribution of $\mathcal{F}$ is

$$\begin{aligned} p(\mathcal{F}|\lambda) &= 2\chi^2(2\mathcal{F}|4,\lambda) && (6.33)\\ &= \left(\frac{2\mathcal{F}}{\lambda}\right)^{1/2} I_1(\sqrt{2\mathcal{F}\lambda})\,\exp\left(-\mathcal{F}-\frac{\lambda}{2}\right) \end{aligned}$$

where I_1 is the modified Bessel function of the first order. In the absence of a signal, this reduces to

$$p(\mathcal{F}|\,0) = \mathcal{F}e^{-\mathcal{F}}\,. \qquad (6.34)$$

We select frequency bins by setting a threshold $\mathcal{F}_{\rm th}$ on the value of the $\mathcal{F}$-statistic in that frequency bin. Given $\mathcal{F}_{\rm th}$, the probabilities for false alarm, false detection and detection are defined analogous to equations (5.6), (5.7) and (5.8):

$$\begin{aligned} \alpha(\mathcal{F}_{\rm th}) &= \int_{\mathcal{F}_{\rm th}}^{\infty} p(\mathcal{F}|\,0)\,d\mathcal{F}\\ &= (1+\mathcal{F}_{\rm th})e^{-\mathcal{F}_{\rm th}}\,, && (6.35)\\ \beta(\mathcal{F}_{\rm th}|\lambda) &= \int_0^{\mathcal{F}_{\rm th}} p(\mathcal{F}|\lambda)\,d\mathcal{F}\,, && (6.36)\\ \eta(\mathcal{F}_{\rm th}|\lambda) &= \int_{\mathcal{F}_{\rm th}}^{\infty} p(\mathcal{F}|\lambda)\,d\mathcal{F}\,. && (6.37) \end{aligned}$$

Note that the relation between α and $\mathcal{F}_{\rm th}$ is different from the relation $\alpha = e^{-\rho_{\rm th}}$ in the non-demodulated case. For small signals, $\eta(\mathcal{F}_{\rm th}|\lambda)$ can be expanded as

$$\eta = \alpha + \frac{\lambda\mathcal{F}_{\rm th}^2}{4}e^{-\mathcal{F}_{\rm th}} + \mathcal{O}(\lambda^2)\,. \qquad (6.38)$$

Once again we will approximate this distribution by a binomial. In fact, we expect the binomial approximation to be better in this as compared to the non-demodulated search because, typically, $T_{\rm coh}$ will now be larger and thus

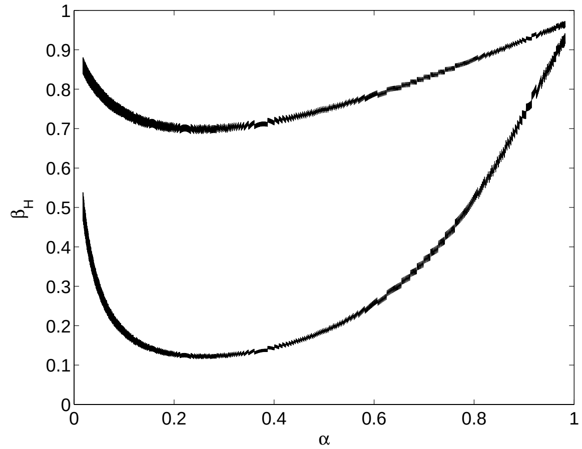

FIG. 16: Graph of β_H as a function of $\alpha = (1+\mathcal{F}_{\rm th})e^{-\mathcal{F}_{\rm th}}$ for $\lambda = 0.2$ (upper curve) and $\lambda = 0.4$ (lower curve). Both curves correspond to $\alpha_H = 0.01$ and $N = 1000$. The minimum of β_H occurs roughly at $\alpha = 0.26$ which corresponds to $\mathcal{F}_{\rm th} = 2.6$.

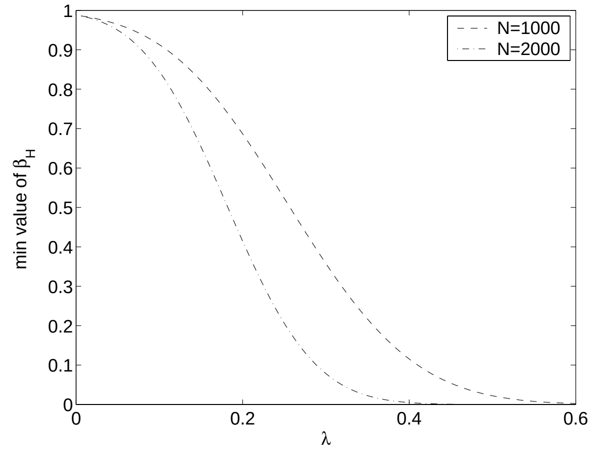

FIG. 17: Graph of the minimum of β_H as a function of λ for $N = 1000$ and 2000. Both curves correspond to $\alpha_H = 0.01$ and $N = 1000$.

the signal will see a more representative 'average' of the detector antenna pattern. Finally, the expressions for the false alarm and false dismissal probabilities in the Hough plane are the same as in equations (5.13) and (5.14) but again with the λ's and η's as above.

With the above definitions at hand, we are now ready to optimize the thresholds $\mathcal{F}_{\rm th}$ and $n_{\rm th}$ using the procedure described in section V B. The differences from that section are simply in the dependence of α on $\mathcal{F}_{\rm th}$ and of η on α. The solution for $n_{\rm th}$ obtained by inverting the equation $\alpha_H(n_{\rm th},\alpha,N)$ given in figure 10 and the analytic approximation of equation (5.22) are unchanged. The graph of β_H as a function of α is however, now different. The result is shown in figure 16. The optimal value for the threshold turns out to be $\mathcal{F}^\star_{\rm th} = 2.6$ corresponding to a false alarm rate of $\alpha^\star = 0.26$. The minimum value of β_H achieved by these thresholds is plotted in figure 17 as a function of λ.

Finally, let us calculate the sensitivity of the search and obtain the analog of equation (5.40). The starting point is again equation (5.31) but now η is related to α by equation (6.38) and α is related to $\mathcal{F}_{\rm th}$ by equation (6.35). Then, ignoring terms of $\mathcal{O}(\lambda^2)$ we get the linear approximation for β_H:

$$\beta_H = \frac{1}{2}\mathrm{erfc}\left(-\mathrm{erfc}^{-1}(2\alpha^\star_H) + \frac{1}{4}\Theta e^{-\mathcal{F}^\star_{\rm th}}(\mathcal{F}^\star_{\rm th})^2\lambda\right) \tag{6.39}$$

where, as before, Θ is given by equation (5.33). Solving for λ in the large N limit leads to

$$\lambda \approx \frac{4\mathcal{S}}{(\mathcal{F}^\star_{\rm th})^2 e^{-\mathcal{F}^\star_{\rm th}}}\sqrt{\frac{2\alpha^\star(1-\alpha^\star)}{N}} \approx \frac{12.73}{\sqrt{N}} \tag{6.40}$$

where, to obtain numerical values, we have taken $\alpha^\star_H = 0.01$, $\beta^\star_H = 0.1$. Now average over the parameters $(\iota,\psi,\alpha,\delta)$ and obtain

$$h_0 = \frac{8.92}{N^{1/4}}\sqrt{\frac{S_n}{T_{\rm coh}}} = 8.92\,N^{1/4}\sqrt{\frac{S_n}{T_{\rm obs}}}\,. \tag{6.41}$$

As expected, this is identical to equation (5.40) except for a slightly different numerical factor. Thus for comparable values of $T_{\rm coh}$ and N, the two versions of the Hough transform search are very similar in sensitivity but the search with with demodulated data does not have any restriction on $T_{\rm coh}$ and will thus lead to a much greater sensitivity, though over a smaller region in parameter space. Thus, if we estimate the astrophysical range of the search as in equation (5.42), we obtain:

$$d = 15.4\,\mathrm{kpc}\times\left(\frac{N}{365}\right)^{\frac{1}{4}}\left(\frac{I_{zz}}{10^{38}\mathrm{kg\text{-}m}^2}\right)\left(\frac{f_r}{500\mathrm{Hz}}\right)^2 \times\left(\frac{\epsilon}{10^{-6}}\right)\left(\frac{T_{\rm coh}}{1\mathrm{day}}\right)^{\frac{1}{2}}\left(\frac{10^{-46}\mathrm{Hz}^{-1}}{S_n}\right)^{\frac{1}{2}}. \tag{6.42}$$

Here we have taken a coherent integration time of 1day and a total observation time of 1yr as the reference values.

VII. CONCLUSIONS

Let us summarize the main ideas and results presented in this work. Since it is not feasible to perform large parameter space searches using the matched filter with presently available computing power, we began by emphasizing the need for hierarchical searches demonstrating the need for an incoherent and computationally inexpensive search method. The Hough transform is an example of such a method. It looks for patterns in the frequency time plane by constructing a histogram in parameter space based on the consistency of observations in the time-frequency plane with an underlying model describing the pattern. We have given a general description of the Hough transform and shown its relevance for pulsar searches.

We have presented two versions of the Hough transform search. The first version takes simple Fourier transforms as input data. This restricts the time baseline of the different segments but it allows us to search over a large sky-patch. The second version takes input data which has been de-modulated to remove the effects of Earth's motion and the spindown of the star; this is achieved by using the $\mathcal{F}$-statistic. We have presented some technical details for both flavors of the search. In particular, we show how to solve the master equation in the two cases and how the use of look-up tables can lead to an enormous saving in computational cost.

We have also analyzed the statistics for both cases and we saw that we need to choose two thresholds: the threshold $\rho_{\rm th}$ or $\mathcal{F}_{\rm th}$ on the coherent statistic used in the two cases, and the threshold $n_{\rm th}$ on the number count in the Hough maps. These thresholds have been chosen in such a way that we get the lowest possible false dismissal rate for a given choice of the false alarm rate. We also estimate the sensitivity of the two flavors of the Hough transform and we find that for the same value of $T_{\rm coh}$ and N, both variations have comparable sensitivity, which improves as $N^{-1/4}T_{\rm coh}^{-1/2}$, as would be expected for an incoherent method that builds on coherent sub-steps. When compared to the sensitivity that a fully coherent search in a very large parameter space would have for the same total observation time $T_{\rm obs}$, the Hough methods are worse by roughly a factor of $N^{1/4}$. Considering that the Hough transform can be expected to run very much faster than any coherent method, it should therefore be able to survey much larger volumes of space than coherent methods, despite its poorer sensitivity in any single direction. This explains why the major gravitational wave collaborations have made this the method of choice for conducting large-scale gravitational wave pulsar surveys.

Acknowledgments

We acknowledge useful discussions with Curt Cutler, Yousuke Itoh, Federico Massaioli, Greg Mendell, and Lucia Zanello. We also acknowledge the support of the Max Planck Society (Albert Einstein Institut) and the Spanish *Ministerio de Ciencia y Tecnología* DGICYT Research Project BFM2001-0988.

APPENDIX A: THE MASTER EQUATION FOR THE SEARCH WITH DEMODULATED DATA

Here we detail the steps leading from equation (6.17) to the master equation (6.18).

Ignoring the amplitude modulation and a possible constant phase, the signal from a pulsar with parameters $(f_{(0)};\vec{\lambda}) = (f_{(0)};\mathbf{n},\{f_{(n)}\})$ would be:

$$h(t,f_{(0)},\vec{\lambda}) = e^{i\Phi(t,f_{(0)},\vec{\lambda})}\,, \tag{A1}$$

where

$$\Phi(t;f_{(0)},\vec{\lambda}) = 2\pi\left[f_{(0)}\Delta t_{\mathbf{n}} + \sum_k \frac{f_{(k)}}{(k+1)!}\left(\Delta t_{\mathbf{n}}\right)^{k+1}\right]\,, \tag{A2}$$

and

$$\Delta t_{\mathbf{n}} = t_{\rm ssb}(t,\mathbf{n}) - t_{\rm ssb}(t_0,\mathbf{n})\,. \tag{A3}$$

Here t_0 is the time in the detector frame to which the frequency and spindown paremeters refer to and $t_{\rm ssb}$ is time in the SSB frame. Neglecting higher order relativistic effects, the detector time t is related to $t_{\rm ssb}$ by

$$t_{\rm ssb}(t,\mathbf{n}) = t + \frac{\mathbf{r}(t)\cdot\mathbf{n}}{c} \tag{A4}$$

where $\mathbf{r}(t)$ is the detector position in the SSB frame.

The DeFT of the pulsar signal (A1) with respect to the demodulation parameters $(f,\vec{\lambda}_d)$ is:

$$\tilde{X}(f) = \int_{-\frac{1}{2}T_{\rm coh}}^{\frac{1}{2}T_{\rm coh}} e^{i[\Phi(t,f_{(0)},\vec{\lambda})-\Phi(t,f,\vec{\lambda}_d)]}dt \tag{A5}$$

Without any loss of generality, we have taken the coherent time interval to be centered arount $t=0$ so that the integral is from $-T_{\rm coh}/2$ to $T_{\rm coh}/2$. Our goal is to determine an analytical expression for the value f^* that maximizes the power $P(f) = \tilde{X}(f)\tilde{X}^*(f)$ in terms of $f_{(0)}$, $\vec{\lambda}_d$ and $\Delta\vec{\lambda}$. This is carried out in two steps: first we Taylor expand around the demodulation parameter $\vec{\lambda}_d$ and then expand around a fiducial time t_1 within the observation interval $(-T_{\rm coh}/2, T_{\rm coh}/2)$.

By Taylor expanding $\Phi(t,f_{(0)},\vec{\lambda})$ around the demodulation parameters $\vec{\lambda}_d = (\mathbf{n}_d, f_{d(k)})$ and retaining terms only upto $\mathcal{O}(1/c)$, we get

$$\begin{aligned}\Delta\Phi(t) &= \Phi(t,f_{(0)},\vec{\lambda}) - \Phi(t,f,\vec{\lambda}_d) \qquad \text{(A6)}\\ &= 2\pi\left[\left(f_{(0)}-f\right)\Delta t_{\mathbf{n}_d} + \sum_{k=1}^{\infty}\frac{\Delta f_{(k)}}{(k+1)!}\left(\Delta t_{\mathbf{n}_d}\right)^{k+1}\right.\\ &\left. + \left\{f_{(0)} + \sum_{k=1}^{\infty}\frac{f_{(k)}}{k!}\left(\Delta t_{\mathbf{n}_d}\right)^{k}\right\}\frac{\Delta\mathbf{r}}{c}\cdot\Delta\mathbf{n} + \mathcal{O}\left(\frac{1}{c^2}\right)\right]\end{aligned}$$

where $\Delta\mathbf{r} = \mathbf{r}(t) - \mathbf{r}(t_0)$, $\Delta\mathbf{n} = \mathbf{n} - \mathbf{n}_d$, and $\Delta f_{(k)} = f_{(k)} - f_{d(k)}$.

Now expand $\Delta\Phi$ about a fiducial time t_1 in the interval $-T_{\rm coh}/2 \le t_1 \le T_{\rm coh}/2$ retaining terms only upto linear order [27]

$$\begin{aligned}\tilde{X}(f) &\approx \int_{-\frac{1}{2}T_{\rm coh}}^{\frac{1}{2}T_{\rm coh}} e^{i\left(\Delta\Phi(t_1)+(t-t_1)\frac{\partial\Delta\Phi}{\partial t}\right)}dt\\ &= e^{i\Delta\Phi(t_1)}\int_{-\frac{1}{2}T_{\rm coh}}^{\frac{1}{2}T_{\rm coh}} e^{i(t-t_1)\frac{\partial\Delta\Phi}{\partial t}}dt\,. \qquad \text{(A7)}\end{aligned}$$

The power $P(f)$ does not depend on $\Delta\Phi(t_1)$, and its maximum is reached for the value of f that satisfies

$$\left.\frac{\partial \Delta\Phi}{\partial t}\right|_{t=t_1} = 0\,. \tag{A8}$$

Differentiating (A6) with respect to t, we get

$$\begin{aligned}
&\frac{1}{2\pi}\left.\frac{\partial \Delta\Phi}{\Delta t}\right|_{t=t_1} \qquad\qquad (A9)\\
&= \left(f_{(0)} - f + \sum_{k=1}^{\infty}\frac{\Delta f_{(k)}}{k!}\,(\Delta t_1)^k\right)\left(1+\frac{\mathbf{v}(t_1)}{c}\cdot\mathbf{n}_d\right)\\
&+ \left(f_{(0)} + \sum_{k=1}^{\infty}\frac{f_{(k)}}{k!}\,(\Delta t_1)^k\right)\frac{\mathbf{v}(t_1)}{c}\cdot\Delta\mathbf{n}\\
&+ \left(\sum_{k=1}^{\infty}\frac{f_{(k)}}{(k-1)!}\,(\Delta t_1)^{k-1}\right)\left(1+\frac{\mathbf{v}(t_1)}{c}\cdot\mathbf{n}_d\right)\frac{\Delta\mathbf{r}_1}{c}\cdot\Delta\mathbf{n}\\
&+\mathcal{O}\left(\frac{1}{c^2}\right)\,,
\end{aligned}$$

where $\Delta t_1 = t_{\rm ssb}(t_1,\mathbf{n}_d) - t_{\rm ssb}(t_0,\mathbf{n}_d)$ and similarly $\Delta\mathbf{r}_1 = \mathbf{r}(t_1) - \mathbf{r}(t_0)$. Setting the right hand side of this equation to zero leads to

$$\begin{aligned}
f - F_0 = &\left(F_0 + \sum_{k=1}^{\infty}\frac{f_{d(k)}}{k!}\,(\Delta t_1)^k\right)\frac{\mathbf{v}(t_1)}{c}\cdot\Delta\mathbf{n} \qquad (A10)\\
&+ \left(\sum_{k=1}^{\infty}\frac{\Delta f_{(k)} + f_{d(k)}}{(k-1)!}\,(\Delta t_1)^{k-1}\right)\frac{\Delta\mathbf{r}_1}{c}\cdot\Delta\mathbf{n} + \mathcal{O}\left(\frac{1}{c^2}\right)\,,
\end{aligned}$$

where F_0 is as given in equation (6.19). Under the assumption the frequency shift can be expressed as a linear combination of the other parameter displacements (as it is the case in the metric approach up to quadratic order), we can remove the dependence on $\Delta f_{(k)}$ in the second line of equation (A10). All the dependence on the residual spindown parameters appears only in the definition of F_0 and we thus get equation (6.18).

---

[1] A. Abramovici et al., Science **256**, 325 (1992).
[2] B. Barish and R. Weiss, Phys. Today **52**, 44 (1999).
[3] B. Willke et al., Class. Quant. Grav., **19(7)**, 1377 (2002).
[4] S. Goßler et al., Class. Quant. Grav., **19(7)**, 1835 (2002).
[5] B. Abbott et al. (The LIGO Scientific Collaboration), Phys. Rev. **D69**, 082004 (2004)
[6] P. Jaranowski, A. Królak, and B.F. Schutz, Phys. Rev. **D58** (1998) 063001.
[7] P.R. Brady, T. Creighton, C. Cutler, and B.F. Schutz, Phys. Rev. **D57** (1998) 2101-2116.
[8] P.R. Brady and T. Creighton, Phys. Rev. **D61** (2000) 082001.
[9] S. Frasca and C La Posta, Il Nuovo Cimento **14 C**, 235 (1991)
[10] P.V.C. Hough, *Machine Analysis of Bubble Chamber Pictures*, International Conference on High Energy Accelerators and Instrumentation, CERN (1959).
[11] P.V.C. Hough, U. S. Patent 3,069,654, 1962.
[12] http://www.lsc-group.phys.uwm.edu/lal/.
[13] C. Bradaschia *et al.*, Nucl. Instrum. Methods **A 289**, 518 (1990).
[14] B. Caron et al., Nucl. Phys. B-Proc. Suppl., **54**, 167 (1997).
[15] B.F. Schutz and M.A. Papa, Proceedings of January 1999 Moriond meeting "Gravitational Waves and Experimental Gravity", `xxx.lanl.gov/abs/gr-qc/9905018`.
[16] M.A. Papa, B.F. Schutz, and A.M. Sintes, ICTP Lecture Notes Series, Volume III (ISBN 92-95003-05-5), p. 431-442 (2001), Italy, Edts. V. Ferrari, J.C. Miller, L. Rezzolla, `xxx.lanl.gov/abs/gr-qc/0011034`.
[17] L. Brocco, S. Frasca, C. Palomba and F. Ricci, Class. Quant. Grav. **20** S655-S664.
[18] P. Astone, S. Frasca, and M.A. Papa, In Proc. of the Aspen Winter Conference on Gravitational Waves and their Detection, ed. by S.Meshkov, 1997.
[19] S. Frasca, Int.J. of Mod. Phys. D **9** 369 (2000)
[20] C. Cutler, I. Gholami, and B. Krishnan, *In preparation.*
[21] B. Allen, M.A. Papa, and B.F. Schutz, Phys. Rev. D **66**, 102003 (2002).
[22] B. Owen, Phys. Rev. **D53**, 6749 (1996).
[23] C. Cutler and K. S. Thorne, In Proceedings of International Conference on General Relativity & Gravitation, ed. by NT Bishop and SD Maharaj, World Scientific, 2002 (`gr-qc/0204090`)

[24] The word hypersurface refers to a sub-manifold of unit co-dimension. The generalization to surfaces of higher co-dimension is straightforward but we shall not discuss it in this paper.
[25] Note that this approximation implies an upper bound on the value of the spindown parameters that we can search over; this will be discussed later in this section in greater detail.
[26] The observation time is now not just a single instant t but is instead a time interval $(t_a, t_a + T_{\rm coh})$. Therefore, instead of using the detector velocity $\mathbf{v}(t)$, we must now use the *average* detector velocity $\bar{\mathbf{v}}(t_a)$ in this time interval. The difference between $\mathbf{v}(t)$ and $\bar{\mathbf{v}}(t_a)$ is not very significant if $T_{\rm coh}$ is less than an hour, as it must be in the non-demodulated case; it is relevant only when $T_{\rm coh}$ becomes comparable to a day, as it will be in the demodulated case.
[27] To see that the choice of t_1 does not matter, note that the sky-patch in which the demodulation is considered valid is such that the power concentrates in a single frequency bin. Therefore, we are free to choose any t_1 value within the interval $(-\frac{1}{2}T_{\rm coh}, \frac{1}{2}T_{\rm coh})$ without affecting the frequency drift.