Marginalizing the likelihood function for modeled gravitational wave searches

Sanjeev Dhurandhar[1,2], Badri Krishnan[2,4] and Joshua L. Willis[3,2]

[1] Inter-University Center for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune 411007, India
[2] Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Callinstr. 38, 30167 Hannover, Germany
[3] Abilene Christian University, Abilene, TX 79699, USA
[4] Leibniz Universität Hannover, D-30167 Hannover, Germany

E-mail: badri.krishnan@aei.mpg.de

Abstract.
Matched filtering is a commonly used technique in gravitational wave searches for signals from compact binary systems and from rapidly rotating neutron stars. A common issue in these searches is dealing with four extrinsic parameters which do not affect the phase evolution of the system: the overall amplitude, initial phase, and two angles determining the overall orientation of the system. The $\mathcal{F}$-statistic maximizes the likelihood function analytically over these parameters, while the $\mathcal{B}$-statistic marginalizes over them. The $\mathcal{B}$-statistic, while potentially more powerful and capable of incorporating astrophysical priors, is not as widely used because of the computational difficulty of performing the marginalization. In this paper we address this difficulty and show how the marginalization can be done analytically by combining the four parameters into a set of complex amplitudes. The results of this paper are applicable to both transient non-precessing binary coalescence events, and to long lived signals from rapidly rotating neutron stars.

PACS numbers: 04.70.-s

1. Introduction

Gravitational wave (GW) searches rely on matched filtering in situations where the expected gravitational wave signal is well known. This is true for both the transient broad-band signals from compact binary coalescence (CBC) events, and for the long lived but narrow-band continuous wave (CW) signals expected from rapidly rotating neutron stars. For both these cases, as we shall explain in more detail later, there are four parameters which do not affect the phase evolution of the system and are independent of the gravitational wave detector(s): the overall amplitude of the signal A, the initial phase of the signal φ_0, the polarization angle ψ, and the inclination angle of the axis of symmetry of the system ι ‡ relative to the line of sight between the system and the detector. There will be other parameters, such as the sky-position, intrinsic parameters such as the masses and spins in a binary system, and the signal frequency and spindown in the case of an isolated neutron star. These parameters may or may-not affect the phase evolution of the signal as seen by a detector, but in any case $(A, \varphi_0, \psi, \iota)$ will not do so. The phase evolution on the other hand, is determined by parameters of arguably greater physical interest such as the masses and spins of the components of a binary system, and in fact $(A, \varphi_0, \psi, \iota)$ are sometimes referred to as "nuisance parameters". We would prefer to not have to explicitly search over $(A, \varphi_0, \psi, \iota)$ and to deal with them analytically as far as possible.

There are two known methods for dealing with such "nuisance" parameters both of which are based on the likelihood function $\Lambda(A, \varphi_0, \psi, \iota|\mathbf{x})$. Here $\mathbf{x}$ is a given data series from a network of GW detectors, and Λ is defined as the ratio

$$\Lambda(A, \varphi_0, \psi, \iota|\mathbf{x}) := \frac{p(\mathbf{x}|A, \varphi_0, \psi, \iota)}{p(\mathbf{x}|\mathbf{noise})} \,. \tag{1}$$

The numerator is the probability density for the data $\mathbf{x}$ for a specific value of the nuisance parameters, while the denominator is the same probability density function in the absence of a signal. We have suppressed the dependence of Λ on other physical source parameters. The first and more commonly used method, is to maximize Λ over the nuisance parameters; this is the standard prescription in the Frequentist framework. This is ultimately motivated by the Neyman-Pearson lemma (see e.g. [1]) which is applicable when we wish to distinguish between two simple hypotheses; here the two hypotheses corresponds to the cases when i) there are no signals present (the null hypothesis), and ii) when we have a signal with fixed values of the nuisance parameters. Consider the problem of finding a region $\mathcal{R}$ in the space of data vectors $\mathbf{x}$ such that the probability of detections is maximized at a fixed false alarm probability. The Neyman-Pearson lemma tells us that level surfaces of the likelihood function yield the most powerful test, i.e. the best choice of the region $\mathcal{R}$ with the largest detection probability for a fixed false alarm probability. For composite hypotheses where all allowed values of the amplitude are included, there is no known analog of the Neyman-Pearson lemma. However, the Neyman-Pearson lemma still motivates us to look at the likelihood function, and to find the values of the nuisance parameters which maximizes it.

In practice for GW searches, one combines the nuisance parameters into quantities (the so-called amplitude parameters) which appear linearly in the GW signal model. This linearity enables one to maximize Λ analytically over the amplitude parameters; in the gravitational wave data analysis context, this was first shown in 1998 [2] for CW signals and soon generalized to the coherent multi-detector case [3], and for CBC searches [4, 5, 6]. The analytic maximization allows one to focus computational resources

‡ The analysis in this paper is restricted to non-precessing systems for which the inclination angle is constant in time.

on numerically maximizing Λ over the other signal parameters determining the phase evolution of the signal.

The second method is to marginalize Λ over the nuisance parameters, i.e. to compute the integral

$$\int \Lambda(A, \varphi_0, \psi, \iota|\mathbf{x})p(A)\, dA\, d\varphi_0\, d\psi\, d\cos\iota\,. \tag{2}$$

Here $p(A)$ is an astrophysical prior on the amplitude. For CBC searches this depends primarily on the assumed spatial distribution of sources, while for CW signals this depends also on properties of neutron stars such as their deformations away from axisymmetry. The volume element corresponds to an isotropic distribution of the angles (φ_0, ψ, ι) which is realistic when we have no other prior information about the system. The Bayesian approach has some benefits such as allowing one to incorporate physical priors, and is potentially more powerful than the $\mathcal{F}$-statistic. Moreover, in the gravitational wave literature it was shown in [7] that if the prior used in the marginalization is in fact the true distribution found in nature, then indeed, the marginalized likelihood function is the optimal detection statistic in the Neyman-Pearson sense (see also Sec. 21.28 of [1]).

However, the above integral is not easy to compute analytically which makes this method less useful in practice. This was first proposed as a detection statistic for CW searches in [8], which defined the so called $\mathcal{B}$-statistic (see also [7]). This work however did not address the critical question of computational cost in evaluating the integral. This was partially addressed in [9] where a different parametrization of the amplitude parameters was constructed based on left- and right-circular polarizations. However useful approximations to the $\mathcal{B}$-statistic integral were found only in the linearly polarized case (where $\cos\iota = 0$). These coordinates were used for CBC searches in [10] which obtained analytic approximations to the integral above, and presented an initial numerical comparison of the Bayesian and Frequentist approaches. Here we propose a new set of complex coordinates (closely related to [9, 10]) for the space of amplitude parameters motivated by properties of the rotation group, which works for both CW and CBC waveforms. These coordinates allow for a more detailed understanding the singularities of the $\mathcal{B}$-statistic and also pave the way for a generalization to include higher modes of the waveform and e.g. precession effects. The same approach also allows for a better understanding of the detector beam pattern functions.

A central message of this paper is the following. (φ_0, ψ, ι) are best thought of as points on a 3-sphere and we can thus identify the 3-sphere with the group of rotations in three dimensional Euclidean space. Given the appropriate tensorial structure of the fields of interest, i.e. fields which transform under rotations as a representation of weight ℓ, we should use the corresponding matrix elements of the rotation group to expand functions of (φ_0, ψ, ι). For gravitational waves, we are dealing with the $\ell = 2$ case and this determines the appropriate basis functions that we should be using to expand functions of (φ_0, ψ, ι). While this may be interesting from a theoretical viewpoint, there is *apriori* no guarantee that this should help simplify the calculation of the $\mathcal{B}$-statistic, but in fact, as we shall show, it does do so.

In the following sections, we start in Sec. 2 with a review of the waveform model, the response of a GW detector and the $\mathcal{F}$- and $\mathcal{B}$-statistics. The new complex amplitudes are defined in Sec.3. Sec. 4 demonstrates how the likelihood function is expressed in terms of the new amplitudes. Sec. 5 shows how Λ can be analytically marginalized over the complex amplitudes along with some example cases and we finally conclude in Sec. 6.

2. Preliminaries and notation

Consider a plane gravitational wave (GW) h_{ab} and an associated right-handed orthonormal wave-frame $(\mathbf{X}, \mathbf{Y}, \mathbf{Z})$ such that the wave is traveling along the $\mathbf{Z}$ direction. In standard linearized general relativity, gravitational waves have two polarizations states defined by the symmetric-tracefree tensors

$$(\mathbf{e}_+)_{ab} = X_a X_b - Y_a Y_b \,, \quad (\mathbf{e}_\times)_{ab} = X_a Y_b + Y_a X_b \,. \tag{3}$$

Then we can write h_{ab} as

$$h_{ab} = h_+ (\mathbf{e}_+)_{ab} + h_\times (\mathbf{e}_\times)_{ab} \,. \tag{4}$$

In this paper we shall use instead the complex null-vector

$$\mathbf{m} = \frac{1}{\sqrt{2}} (\mathbf{X} + i\mathbf{Y}) \,. \tag{5}$$

It is easy to see that $\mathbf{m} \cdot \mathbf{m} = 0$ and $\mathbf{m} \cdot \mathbf{m}^\star = 1$. In general the gravitational wave h_{ab} can then be written as

$$h_{ab} = \mathfrak{h}^\star m_a m_b + \mathfrak{h} m^\star_a m^\star_b \,, \tag{6}$$

where the complex scalar $\mathfrak{h}$ contains the two polarizations $h_{+,\times}$:

$$\mathfrak{h} = h_+ + i h_\times \,. \tag{7}$$

Using m_a instead of X_a and Y_a is equivalent to using left and right circular polarizations instead of linear polarization states. Note that if we perform a counter-clockwise rotation in the $(\mathbf{X}, \mathbf{Y})$ plane by an angle ψ:

$$\mathbf{X} \to \ \cos\psi \mathbf{X} + \sin\psi \mathbf{Y} \,, \tag{8}$$

$$\mathbf{Y} \to \ -\sin\psi \mathbf{X} + \cos\psi \mathbf{Y} \,, \tag{9}$$

then $\mathbf{m} \to e^{-i\psi}\mathbf{m}$, and thus $\mathfrak{h} \to e^{-2i\psi}\mathfrak{h}$; m_a is, by definition, assigned a spin weight $+1$, and $\mathfrak{h} = m^a m^b h_{ab}$ is said to have spin-weight $+2$. For an elliptically polarized wave, it is always possible to choose the wave frame aligned with the principal directions of the polarization ellipse. In this case h_+ and $h_\times$ have a $\pi/2$ offset in phase:

$$h_+(t) = A_+(t) \cos\Phi(t) \,, \qquad h_\times(t) = A_\times(t) \sin\Phi(t) \,. \tag{10}$$

Here the amplitudes $A_{+,\times}$ are slowly varying functions of time while the phase $\Phi(t)$ is rapidly varying. We shall assume that the wave-frame is aligned with the principal polarization directions in this way.

Consider now an interferometric GW detector on Earth. We shall work in the long-wavelength approximation (appropriate for current ground based detectors) where the GW wavelength λ is much larger than the interferometer arm-length. In this approximation, the strain $h(t)$ measured by the detector will be given as

$$h(t) = h_{ab}D^{ab} \tag{11}$$

where D_{ab} is the detector tensor

$$D_{ab} = \frac{1}{2}(u_a u_b - v_a v_b)\,. \tag{12}$$

(we refer the reader to e.g. [11] for further details, and to e.g. [12, 13] when it is necessary to go beyond the long-wavelength approximation). Here $\mathbf{u}, \mathbf{v}$ are unit-vectors along the detector arms, and we follow the convention that $(\mathbf{u}, \mathbf{v}, \mathbf{z})$ forms a right handed coordinate system with $\mathbf{z}$ pointing away from Earth's center. Let us also take the angle between $\mathbf{u}$ and $\mathbf{v}$ to be 2ζ. Take a frame $(\mathbf{x}, \mathbf{y})$ such that $\mathbf{x}$ bisects the arms of the detector. In this case $D_{ab} = \frac{1}{2}\sin 2\zeta(x_a y_b + y_a x_b)$. Unless mentioned otherwise, we henceforth take the arms to be perpendicular and keep in mind that an overall factor of $\sin 2\zeta$ suffices to account for it if necessary.

Introduce spherical polar coordinates (θ, ϕ) associated with the detector frame $(\mathbf{x}, \mathbf{y}, \mathbf{z})$ in the usual way, and the corresponding unit vectors $(\mathbf{e}_\theta, \mathbf{e}_\phi)$. The direction of propagation of the GW, i.e. $\mathbf{Z}$ is pointed towards the detector. We can take $(\mathbf{e}_\theta, -\mathbf{e}_\phi, \mathbf{Z})$ as a reference wave frame and $\mathbf{m}^{(0)} = (\mathbf{e}_\theta - i\mathbf{e}_\phi)/\sqrt{2}$ as a reference null spin vector. The wave-frame (X, Y) will be related to $(-\mathbf{e}_\theta, \mathbf{e}_\varphi)$ by a counter-clockwise rotation (as determined by $\mathbf{Z}$ being inward pointing). This implies that we can rotate $\mathbf{m}^{(0)}$ and align it with $\mathbf{m}$: $\mathbf{m} = e^{-i\psi}\mathbf{m}^{(0)}$. With these conventions, we can rewrite h_{ab} from Eq. (6) as

$$h_{ab} = \mathfrak{h}^\star e^{-2i\psi} m_a^{(0)} m_b^{(0)} + \mathfrak{h} e^{2i\psi} m_a^{(0)\star} m_b^{(0)\star}\,. \tag{13}$$

The detector response can be written in terms of a complex beam pattern function $\mathfrak{F}(\theta, \phi, \psi)$ as

$$h(t) = \frac{1}{2}\mathfrak{h}\mathfrak{F}^\star(\theta, \phi, \psi) + \frac{1}{2}\mathfrak{h}^\star\mathfrak{F}(\theta, \phi, \psi) \tag{14}$$

with $\mathfrak{F} = F_+ + iF_\times$. The factor of $1/2$ has been chosen so that the above is equivalent to $h = F_+ h_+ + F_\times h_\times$. From Eqs. (13) and (11) we get

$$\mathfrak{F}(\theta, \phi, \psi) = 2D^{ab} m_a m_b = 2e^{-2i\psi} D^{ab} m_a^{(0)} m_b^{(0)} \tag{15}$$

$$= \left(\frac{1 + \cos^2\theta}{2}\sin 2\phi - i\cos\theta\cos 2\phi\right) e^{-2i\psi}\,. \tag{16}$$

We often write Eq. (15) as

$$\mathfrak{F}(\theta, \phi, \psi) = \mathfrak{f}(\theta, \phi) e^{-2i\psi}\,. \tag{17}$$

This separates the polarization angle ψ from the sky-position. As in [2], it is useful to define the ψ independent beam pattern functions $a(\theta, \phi)$ and $b(\theta, \phi)$ as

$$\mathfrak{f}(\theta, \phi) = a(\theta, \phi) + ib(\theta, \phi)\,. \tag{18}$$

This leads to

$$F_+ = a\cos 2\psi + b\sin 2\psi\,,\quad F_\times = b\cos 2\psi - a\sin 2\psi\,. \tag{19}$$

When dealing with multiple detectors, one refers (θ, ϕ, ψ) to a common coordinate system such a geocentric or a solar system barycenter based system. In this case we can still separate the sky-positions from the polarization angle.

The time parameter t appearing in Eq. (14) is the arrival time of wavefronts at the detector. It is related to the time of arrival at the origin of coordinates τ according to $t = \tau - \mathbf{r}\cdot\mathbf{n}/c$ where $\mathbf{r}$ is the position of the detector, and $\mathbf{n}$ the direction to the source and c the speed of light. For transient signals it is common to use a geocentric system while for CW sources a solar-system barycentric system is more natural. For long duration signals this leads to the periodic Doppler shifting of the signal frequency as Earth rotates around its axis and orbits the Sun, and furthermore, it is often necessary to take into account the relativistic Einstein and Shapiro time delays [2].

2.1. The relation between the beam pattern functions and representations of the rotation group

It is interesting to note that, as first done in [14], $\mathfrak{F}$ can be written in the following way:

$$\mathfrak{F} = ie^{-2i\phi}\frac{(1-\cos\theta)^2}{4}e^{-2i\psi} - ie^{2i\phi}\frac{(1+\cos\theta)^2}{4}e^{-2i\psi}\,. \tag{20}$$

As shown in [14] the two terms appearing here are essentially matrix elements of the rotation group, which in turn are related to the spin-weighted spherical harmonics [15, 16] (see Appendix A). It is in fact generally true that the detector response function can be expanded in terms of the matrix elements of the rotation group. To see this we briefly review some results from the representation theory of the rotation group in 3-dimensional Euclidean space. We later apply the same approach to the emitted GW signal emitted by a CBC or CW source. We shall mostly follow [17].

Let $\mathcal{G}$ denote the group of rotations in $\mathbb{R}^3$ and let $g \in \mathcal{G}$ be a particular rotation. Let the matrix elements of g be g_{ij} (satisfying $\mathbf{g}^\intercal\mathbf{g} = 1$) which transforms a vector x_i as $x'_i = \sum_{j=1}^{3} g_{ij}x_j$. It will be convenient to use the Euler angles to represent a general rotation. We use the 'zxz' convention where a general rotation is composed of three rotations: i) a rotation around the z-axis by an angle φ, ii) a rotation by an angle θ around the x-axis, and finally, iii) a rotation by an angle ψ around the z-axis. Thus, we write $g(\varphi, \theta, \psi)$ for a general rotation. The inverse of $g(\varphi, \theta, \psi)$ is seen to be $g(\pi - \psi, \theta, \pi - \varphi)$. We take all transformations to be active.

Consider now an irreducible representation of $\mathcal{G}$ of weight ℓ. Thus, we have a $2\ell + 1$ dimensional complex vector space, and each rotation g is represented by a $2\ell + 1$ dimensional unitary matrix $T^\ell_{mn}(g)$, with $-\ell \le m, n \le \ell$. We shall often drop the superscript ℓ when it is obvious what the weight is, and we shall also write $T^\ell_{mn}(\varphi, \theta, \psi)$, i.e. as functions on $\mathcal{G}$. The product of two rotations is written as a matrix multiplication

$$T^\ell_{mn}(g_1g_2) = \sum_{k=-\ell}^{\ell} T^\ell_{mk}(g_1)T^\ell_{kn}(g_2)\,. \tag{21}$$

This equation can be viewed in the following way: under the transformation $g \to gg_1$, the rotation group acts on the matrix elements themselves. Thus, each row of the matrix T^ℓ_{mn} (i.e. for fixed m) transforms as an irreducible representation of the rotation group. The general expression for the matrix elements are summarized in Appendix A.

An important fact we shall use is the following: The functions $T^\ell_{mn}(\varphi, \theta, \psi)$ form a complete orthogonal basis in the Hilbert space of square integrable functions $f(\varphi, \theta, \psi)$ on $\mathcal{G}$. This fact can be used to expand a general tensor on the 2-sphere S^2 (such as a gravitational wave arriving from different sky-directions and with different polarization angles) by the following construction. Start with the standard triad at the north-pole i.e. the basis vectors along the x, y, z directions. Let us denote this triad as $\mathbf{e}_i$ $(i = 1, 2, 3)$ and consider a gravitational wave arriving from a sky-position P with coordinates (θ, ϕ), and with a polarization angle ψ. The rotation $g(\pi, \theta, \phi+\pi/2)$ takes the north pole to the point (θ, ϕ) and the triad at the north-pole is transformed to $(-\mathbf{e}_\phi, \mathbf{e}_\theta, \mathbf{e}_r)$. The inverse rotation $g(\pi/2 - \phi, \theta, 0)$ takes the point P to the north pole and takes the wave-frame with $\psi = 0$ to the triad at the north pole. To go to the wave-frame with a polarization angle ψ, the transformation is $g(\pi - \psi, \theta, \phi + \pi/2)$. This construction yields a triad for every group element, i.e. for all values of (φ, θ, ψ). By definition then, each of the basis vectors then transforms as a scalar on $\mathcal{G}$ under the transformation $g \to gg_1$. The components of any tensor field in this basis are then to be viewed as functions of (φ, θ, ψ) and thus can be expanded in terms of T^ℓ_{mn}. An important ingredient here is the spin-weight: any quantity with spin weight s can be expanded in terms of T^ℓ_{sm}. A scalar corresponds to $s = 0$, and in fact the functions T^ℓ_{0m} are just the usual spherical harmonics $Y_{\ell m}$. This is just the familiar observation that any scalar function can be expanded in terms of the spherical harmonics. For gravitational waves, the relevant tensorial structure is that of symmetric trace-free tensors and, we shall see that the appropriate basis functions are $T^\ell_{-2,m}$.

Let us now return to the beam pattern function and write it in terms of the matrix elements which are listed in Appendix A. The relevant matrix elements of this transformation are

$$T^\ell_{mn}(\pi - \psi, \theta, \phi + \pi/2) = (-1)^{m+n} i^n e^{im\psi} P^\ell_{mn}(\cos\theta) e^{-in\phi} \,. \tag{22}$$

The complex antenna pattern of Eq. (20) can be written as

$$\mathfrak{F}(\theta, \phi, \psi) = iT^2_{-2,-2}(\pi - \psi, \theta, \phi + \pi/2) - iT^2_{-2,2}(\pi - \psi, \theta, \phi + \pi/2) \,. \tag{23}$$

If we go back to the more general expression Eq. (17) for the beam pattern function, we see that $\mathfrak{F}$ is best viewed as a function on $\mathcal{G}$ or, equivalently, the 3-sphere S^3. In addition, $\mathfrak{F}$ is seen to have spin weight -2 because of the factor of $e^{2i\psi}$. This implies that in general $\mathfrak{F}$ can be expanded in terms of $T^\ell_{-2,m}$. Typically, only the $\ell = 2$ components are considered in the literature. There are some exceptions such as [18] which considers the $\ell = 3$ contributions.

As an example of the utility of this formalism, consider a situation when we use a coordinate system not centered on the detector. For a detector on Earth, it is useful to consider a geocentric coordinate system. Thus, let (Θ_I, Φ_I) be the position of an interferometric detector, and let Ψ_I be the orientation of its arms with respect to the lines of latitude. Let g be the transformation from the geocentric system to the wave frame, and let g_I be the transformation from the geocentric system to the detector frame. Then, the transformation $g * g_I^{-1}$ takes us from the detector to the wave-frame. From Eq. (21) we get for example

$$T^2_{-2,-2}(gg_I^{-1}) = \sum_{s=-2}^{2} T^2_{-2,s}(g) T^2_{s,-2}(g_I^{-1}) \tag{24}$$

$$= \sum_{s=-2}^{2} T^2_{-2,s}(\pi - \psi, \theta, \phi + \pi/2) T^2_{s,-2}(\pi/2 - \Phi_I, \Theta_I, \Psi_I)\,. \tag{25}$$

Viewed as an expansion on the sky, this is an expansion in the five dimensional space spanned by $T^2_{-2,s}$ for $-2 \leq s \leq 2$ and the coefficients of the expansion depend on the location and orientation of the detector. Using the expressions in Appendix A, it is straightforward to obtain analytic expressions for these coefficients and thus for the beam pattern functions.

2.2. The $\mathcal{F}$- and $\mathcal{B}$-statistics

We now summarize the conventional approach to matched filtering in GW data analysis using the so-called $\mathcal{F}$ and $\mathcal{B}$ statistics. Let us consider M detectors labeled by an index $I = 1, 2 \ldots M$. As in Eq. (10) taking $\Phi(t) = 2\varphi_0 + 2\varphi(t)$, we write $h_{+,\times}$ in terms of a slowly varying amplitude $A(t)$ and a rapidly varying phase $\varphi(t)$:

$$h_+ = A_+(t)\cos(2\varphi_0 + 2\varphi(t))\,,$$
$$h_\times = A_\times(t)\sin(2\varphi_0 + 2\varphi(t))\,. \tag{26}$$

Here we include a factor of 2 with the phase as is natural for GW sources, and φ_0 is an initial phase which depends on a choice of frame attached with the source. For CBC systems, $\varphi(t)$ is the orbital phase of the binary system while in the CW case this is the rotational phase of the neutron star. For the dominant $\ell = m = 2$ mode for non-precessing systems, including both CW and CBC sources, we write the amplitudes $A_{+,\times}$ as

$$A_+ = A\eta(t)\frac{1+\cos^2\iota}{2}\,, \qquad A_\times = A\eta(t)\cos\iota\,. \tag{27}$$

Here $\eta(t)$ is a slowly varying function of time, ι is the angle between the line of sight to the system and the system's axis of rotation; for non-precessing systems ι would not be time dependent.

For CW sources, $A_{+,\times}(t)$ are in fact constants over time and are written in terms of an overall amplitude h_0. Thus, $A = h_0$ and $\eta(t) = 1$. The amplitude h_0 depends inversely on the distance and additionally it depends on physical properties of say the neutron star crust, the fluid motion in the interior etc. For the non-precessing CBC case, we have the same dependence on the angle ι. In addition, the amplitude will depend on the masses and spins of the binary components and the distance D to the binary. Furthermore, η is no longer constant, but instead will increase as the separation between the binary components decreases (see e.g. [19]). It is convenient, as in [5], to write this as $A = D_0/D$ where D_0 is a fiducial distance.

We can write h_+ and $h_\times$ as

$$\begin{aligned} h_+ &= A_+(t)\cos 2\varphi_0 \cos 2\varphi(t) - A_+(t)\sin 2\varphi_0 \sin 2\varphi(t)\,, \\ h_\times &= A_\times(t)\sin 2\varphi_0 \cos 2\varphi(t) + A_+(t)\cos 2\varphi_0 \sin 2\varphi(t)\,. \end{aligned} \tag{28}$$

Combining $h_{+,\times}$ and the beam pattern functions $F_{+,\times}$, and separating out the polarization angle explicitly, it is easy to show that the signal in detector I can be written as

$$h^I(t) = \sum_{\mu=1}^{4} \mathcal{A}^\mu h^I_\mu(t)\,, \tag{29}$$

where $\mathcal{A}^\mu,\ \mu = 1, 2, 3, 4$ are amplitudes depending on the distance to the source, initial phase φ_0, polarization angle ψ and the inclination angle ι of the source and are given by:

$$\begin{aligned} \mathcal{A}^1 &= A_+ \cos 2\varphi_0 \cos 2\psi - A_\times \sin 2\varphi_0 \sin 2\psi\,, \\ \mathcal{A}^2 &= A_+ \cos 2\varphi_0 \sin 2\psi + A_\times \sin 2\varphi_0 \cos 2\psi\,, \\ \mathcal{A}^3 &= -\,A_+ \sin 2\varphi_0 \cos 2\psi - A_\times \cos 2\varphi_0 \sin 2\psi\,, \\ \mathcal{A}^4 &= -\,A_+ \sin 2\varphi_0 \sin 2\psi + A_\times \cos 2\varphi_0 \cos 2\psi\,, \end{aligned} \tag{30}$$

It is important to note that the $\mathcal{A}^\mu$ are detector independent; for multiple detectors they are defined in a common coordinate system (geocentric or solar system barycenter). The detector dependent basis signals $h^I_\mu(t)$ are defined by,

$$h^I_1(t) = a^I \eta(t) \cos 2\varphi(t_I)\,, \qquad h^I_2 = b^I \eta(t) \cos 2\varphi(t_I)\,, \tag{31}$$

$$h^I_3(t) = a^I \eta(t) \sin 2\varphi(t_I)\,, \qquad h^I_4 = b^I \eta(t) \sin 2\varphi(t_I)\,. \tag{32}$$

where t_I is the retarded time in detector I. It is conventional in the CBC literature to define $h_o(t) = \eta(t)\cos 2\varphi(t)$ and $h_{\pi/2} = \eta(t)\sin 2\varphi(t)$.

The data in detector I with signal present is:

$$x^I(t) = h^I(t) + n^I(t)\,, \tag{33}$$

where $n^I(t)$ is the noise in detector I. The multi-detector data vector is then $\mathbf{x}(t)$. A scalar product can be defined for each detector I on two data trains x^I and y^I as,

$$(x^I|y^I)_I = 4\Re\left(\int_0^\infty df\, \frac{\tilde{x}^I(f)\tilde{y}^{I*}(f)}{S^I_n(f)}\right)\,, \tag{34}$$

where $S^I_n(f)$ is the one sided PSD of the noise in detector I. If the noises between detectors are uncorrelated then a useful multi-detector scalar product between vectors $\mathbf{x}$ and $\mathbf{y}$ can be defined for the network as:

$$(\mathbf{x}|\mathbf{y}) = \sum_I (x^I|y^I)_I\,. \tag{35}$$

We do not attach any subscript to the network scalar product - it may be understood from the context. The multi-detector log-likelihood is then given by:

$$\ln\Lambda(\mathcal{A}^\mu|\mathbf{x}) = (\mathbf{x}|\mathbf{h}) - \frac{1}{2}(\mathbf{h}|\mathbf{h}) = \mathcal{A}^\mu x_\mu - \frac{1}{2}\mathcal{A}^\mu\mathcal{M}_{\mu\nu}\mathcal{A}^\nu\,, \tag{36}$$

where

$$x_\mu = (\mathbf{x}|\mathbf{h}_\mu)\,, \qquad \text{and} \qquad \mathcal{M}_{\mu\nu} = (\mathbf{h}_\mu|\mathbf{h}_\nu)\,. \tag{37}$$

The log-likelihood then may be maximized with respect to the amplitudes $\mathcal{A}^\mu$ to obtain the maximum likelihood or the $\mathcal{F}$ statistic:

$$\mathcal{F} \equiv [\ln\Lambda(\mathbf{x})]_{\max} = \frac{1}{2}x_\mu\mathcal{M}^{\mu\nu}x_\nu\,, \tag{38}$$

where $\mathcal{M}^{\mu\nu}$ is the inverse of $\mathcal{M}_{\mu\nu}$ when the inverse exists - that is when $\mathcal{M}$ is non-singular. The usual coherent multi-detector statistic is the coherent signal to noise ratio (SNR) which is just $2\mathcal{F}$. That we have a maximum of Λ rather than an extremum requires that the eigenvalues of $M_{\mu\nu}$ be positive definite. We shall discuss this explicitly later.

The $\mathcal{M}$ matrix consists of the antenna pattern functions and the noise variances. For the first matrix element, consider first the CBC case. Here the duration of the signal can be considered short enough that a^I and b^I are, to a very good approximation, constant in time. This approximation starts breaking down only when the duration cannot be considered to be much smaller than a sidereal day. Also, it is not difficult to see that $(h_0^I|h_0^I) \approx (h_{\pi/2}^I|h_{\pi/2}^I)$, and $(h_0^I|h_{\pi/2}^I) \approx 0$. Thus, the structure of this matrix (in the long-wavelength limit) is:

$$\mathcal{M} = \begin{bmatrix} A & C & 0 & 0 \\ C & B & 0 & 0 \\ 0 & 0 & A & C \\ 0 & 0 & C & B \end{bmatrix}\,, \tag{39}$$

where $A = \mathbf{a}\cdot\mathbf{a}, B = \mathbf{b}\cdot\mathbf{b}$ and $C = \mathbf{a}\cdot\mathbf{b}$, where $\mathbf{a}$ is the M dimensional vector whose components are $a^I,\ I = 1, 2, ..., M$ and $\mathbf{b}$ is defined similarly. The dot product is weighted by the noise variance in each detector. For example, $\mathbf{a}\cdot\mathbf{b} := \sum_I \sigma_I^2 a^I b^I$ where $\sigma_I^2 := (h_0^I|h_0^I) = (h_{\pi/2}^I|h_{\pi/2}^I)$.

For CW waveforms, $\eta(t)$ is a constant, and the signals are narrow band. If the signal frequency is f_0, and in the time domain the data duration is T_0 centered at the origin, then using Parseval's inequality and that $\sin^2(2\varphi(t))$ averages to 1/2 over many cycles, we get for example

$$(h_1^I|h_1^I)_I \approx \frac{2}{S_n^I(f_0)}\int_{-\infty}^{\infty}\widetilde{h}_1^{I\star}(f)\widetilde{h}_1^I(f)df \approx \frac{1}{S_n^I(f_0)}\int_{-T_0/2}^{T_0/2}(a^I(t))^2dt\,. \tag{40}$$

Thus,

$$\mathcal{M}_{11} = \sum_I \frac{1}{S_n^I(f_0)}\int_{-T_0/2}^{T_0/2}(a^I(t))^2dt\,. \tag{41}$$

It is easy to see that $\mathcal{M}_{\mu\nu}$ has the same form as Eq. (39). The matrix elements can be written in terms of integrals as above, which in turn can again be expressed as an inner product as for the CBC case.

The $\mathcal{B}$ statistic defined in [8] is a Bayesian statistic which employs a prior probability distribution on the amplitudes. Here Λ is a functional of the data $\mathbf{x}$ and depends only on the amplitude parameters $\mathcal{A}^\mu$, that is, $\Lambda(\mathbf{x};\mathcal{A}^\mu)$ and is given by the expression in Eq. (36). There is also a prior on the amplitudes which must be supplied and we denote this by $p_{\mathcal{A}}(\mathcal{A}^\mu)$ and is a function of the 4 amplitudes $\mathcal{A}^\mu$. Then the $\mathcal{B}$ statistic is defined as:

$$\mathcal{B}(\mathbf{x}) = \int \Lambda(\mathbf{x};\mathcal{A}^\mu) p_{\mathcal{A}}(\mathcal{A}^\mu) d^4\mathcal{A}\,. \tag{42}$$

It was further shown in [8] that a uniform amplitude prior on the $\mathcal{A}^\mu$, that is, $p_{\mathcal{A}}(\mathcal{A}^\mu) =$ const. leads one back to the $\mathcal{F}$ statistics for a targeted search, that is, the $\mathcal{B}$ statistic is equivalent to the $\mathcal{F}$ statistic. As done in [8], we can assign a uniform prior on the physical parameters $A, \cos\iota, \varphi_0, \psi$, where one then assumes that $0 \leq A \leq A_{\rm max}$. Then a Jacobian factor enters into the integral in Eq. (42), which is in fact the physical prior. The Jacobian factor is:

$$\mathcal{J} = \frac{\partial(A^1, A^2, A^3, A^4)}{\partial(A, \cos\iota, \varphi_0, \psi)} = \frac{1}{2} A^3 (1 - \cos^2\iota)^3\,. \tag{43}$$

and then the integrand is $\Lambda(\mathbf{x};\mathcal{A}^\mu)/\mathcal{J}$, where the integration variables are $\mathcal{A}^\mu$. The implicit prior in the $\mathcal{F}$ statistic is just $\mathcal{J}$.

3. New complex amplitudes

The standard amplitudes $\mathcal{A}^\mu$ defined in Eq. (30) are not convenient for computing the $\mathcal{B}$-statistic. For this reason a new set of amplitudes, linear combinations of the $\mathcal{A}^\mu$, were defined in [9, 10]. These amplitudes may be viewed as being based on left- and right-circularly polarized signals rather than linearly polarized signals. We shall below define a new set of amplitudes which are complexified versions of the amplitudes defined in [9, 10]. This discussion will cover both CW and CBC signals.

We derive these alternate complex amplitude parameters starting with the complex strain $\mathfrak{h} = h_+ + i h_\times$, and setting $\cos\iota = \chi$:

$$\mathfrak{h} = A\left(\frac{1+\chi^2}{2}\right)\eta(t)\cos(2\varphi_0 + 2\varphi(t)) + iA\chi\eta(t)\sin(2\varphi_0 + 2\varphi(t)) \tag{44}$$

$$= Ae^{2i\varphi_0}\frac{(1+\chi)^2}{4}\mathfrak{h}_0(t) + Ae^{-2i\varphi_0}\frac{(1-\chi)^2}{4}\mathfrak{h}_0^\star(t) \tag{45}$$

where

$$\mathfrak{h}_0(t) = \eta(t) e^{2i\varphi(t)}\,. \tag{46}$$

Combining this with the complex beam pattern function according to Eq.(14) yields

$$h^I = \sum_{\mu=1}^{4} \mathcal{B}^\mu \mathfrak{h}_\mu^I \tag{47}$$

where

$$\mathcal{B}_1 = A\, e^{-2i\varphi_0}\, \frac{(1+\chi)^2}{4}\, e^{-2i\psi}\,, \qquad \mathcal{B}_2 = A\, e^{-2i\varphi_0}\, \frac{(1-\chi)^2}{4}\, e^{2i\psi}\,,$$
$$\mathcal{B}_3 = A\, e^{2i\varphi_0}\, \frac{(1-\chi)^2}{4}\, e^{-2i\psi}\,, \qquad \mathcal{B}_4 = A\, e^{2i\varphi_0}\, \frac{(1+\chi)^2}{4}\, e^{2i\psi}\,. \tag{48}$$

The detector dependent complex basis functions $\mathfrak{h}^I_\mu$ are obtained by using the equations Eq. (29) and Eq. (47). These complex basis functions are:

$$\mathfrak{h}^I_1 = \mathfrak{h}^{I\star}_4 = \frac{1}{2}\mathfrak{f}^I \mathfrak{h}^{I\star}_0\,, \qquad \mathfrak{h}^I_2 = \mathfrak{h}^{I\star}_3 = \frac{1}{2}\mathfrak{f}^{I\star} \mathfrak{h}^{I\star}_0\,. \tag{49}$$

Since $\mathcal{B}_3 = \mathcal{B}_2^\star$ and $\mathcal{B}_4 = \mathcal{B}_1^\star$, we could work just with $\mathcal{B}_2$ and $\mathcal{B}_2$. It shall however be convenient to consider all the four $\mathcal{B}$'s when discussing the transformation from the four real amplitudes. As noted earlier for the beam pattern function, these are related to the rotation matrix group elements, and this is a reflection of the fact that, just as for the detector response function, one should view φ_0, ψ, χ as coordinates on the group of rotations. The relations above can be easily inverted

$$\chi = \frac{\sqrt{|\mathcal{B}_1|} - \sqrt{|\mathcal{B}_2|}}{\sqrt{|\mathcal{B}_1|} + \sqrt{|\mathcal{B}_2|}}\,, \qquad \sqrt{A} = \sqrt{|\mathcal{B}_1|} + \sqrt{|\mathcal{B}_2|} \tag{50}$$
$$\psi = -\frac{1}{4}\text{arg}\left(\frac{\mathcal{B}_1}{\mathcal{B}_2}\right)\,, \qquad \psi = -\frac{1}{4}\text{arg}\left(\frac{\mathcal{B}_1}{\mathcal{B}_2^\star}\right)\,. \tag{51}$$

One can easily verify that the $\mathcal{B}_\mu$ are linear combinations of $\mathcal{A}_\mu$ and vice-versa. Writing the $\mathcal{B}_\mu$ and $\mathcal{A}_\mu$ as column vectors $\mathbf{B}$ and $\mathbf{A}$, we find that $\mathbf{A} = \mathbf{S}\mathbf{B}$, where $\mathbf{S}$ is a 4×4 matrix given by:

$$\mathbf{S} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 1 & 1 \\ i & -i & i & -i \\ -i & -i & i & i \\ 1 & -1 & -1 & 1 \end{bmatrix}\,. \tag{52}$$

The determinant of $\mathbf{S}$ is unity and $\mathbf{S}$ is unitary so that $\mathbf{S}^\dagger\mathbf{S} = I$, where I is a 4×4 unit matrix; $\mathbf{S}^\dagger = (\mathbf{S}^\intercal)^*$. Thus we have $\mathbf{B} = \mathbf{S}^\dagger\mathbf{A}$.

We further note that $\mathcal{B}_4 = \mathcal{B}_1^*$ and $\mathcal{B}_3 = \mathcal{B}_2^*$. This fact is useful because when we finally write the integral for the $\mathcal{B}$ statistic. The volume element $d\mathcal{B}_1 d\mathcal{B}_2 d\mathcal{B}_3 d\mathcal{B}_4 = d\mathcal{B}_1 d\mathcal{B}_1^* d\mathcal{B}_2 d\mathcal{B}_2^*$, which is then a product of two area elements. If we assume again a uniform prior in the physical variables A, χ, φ_0, ψ, we will also need the Jacobian in terms of the new amplitude variables $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3, \mathcal{B}_4$. In fact the Jacobian is unchanged because $\mathbf{S}$ is unitary and we have the result:

$$\mathcal{J} = \frac{\partial(\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3, \mathcal{B}_4)}{\partial(A, \chi, \varphi_0, \psi)} = \frac{1}{2}A^3(1-\chi^2)^3 \equiv 32(\mathcal{B}_1\mathcal{B}_2\mathcal{B}_3\mathcal{B}_4)^{3/4}\,. \tag{53}$$

The Jacobian in terms of the new amplitudes is remarkably just a product of the $\mathcal{B}_\mu$.

4. Likelihood in terms of the complex amplitudes

We begin with the definition of a suitable inner-product for two complex signals $x(t)$ and $y(t)$ in a detector labeled by the index I; $\widetilde{x}(f)$ and $\widetilde{y}(f)$ will denote their Fourier transforms. The usual definition of Eq. (35) works only for real signals. Let $S^I(f)$ now be the *double-sided* power-spectral density of the detector noise. For real-valued signals $x(t)$, for which $\widetilde{x}^\star(f) = \widetilde{x}(-f)$ it is conventional to only focus on positive frequencies and work instead with the single-sided power spectral density. However, since we need to deal with complex functions, both positive and negative frequencies need to be considered. The inner product of $(x|y)$ of x and y is then defined as

$$(x|y)_I := \int_{-\infty}^{\infty} \frac{\widetilde{x}^\star(f)\widetilde{y}(f)}{S_n^I(f)} df \,. \tag{54}$$

Note that this is a complex inner product so that, for example, $(x|y)_I = (y|x)_I^\star$ and $(c_1 y_1 + c_2 y_2 | x)_I = c_1^\star (y_1|x)_I + c_2^\star (y_2|x)_I$. Here $c_{1,2}$ are complex numbers and $x, y, y_{1,2}$ are complex time series. Note that the inner product is specific for a particular detector through its noise spectral density. Consider a collection of M detectors with uncorrelated noises labeled by $I = 1 \ldots M$. Let $x_I(t)$ be time series in each of the detectors which will collectively be denoted in boldface $\mathbf{x}$. Then, the multi-detector inner product is

$$(\mathbf{x}|\mathbf{y}) := \sum_{I=1}^{M} (x_I|y_I)_I = \sum_{I=1}^{M} \int_{-\infty}^{\infty} \frac{\widetilde{x}_I^\star(f)\widetilde{y}_I(f)}{S_n^I(f)} df \,. \tag{55}$$

For real signals, this inner product is equivalent to Eq. (35), which is why we use the same notation for both inner products.

In terms of the $\mathcal{B}^\mu$, and the signal decomposition of Eq. (47) it is apparent that the log-likelihood function is still quadratic:

$$\ln \Lambda = \sum_{\mu=1}^{4} \mathcal{B}^\mu (\mathbf{x}|\mathfrak{h}^\mu) - \frac{1}{2}\mathcal{B}^{\mu\star}\mathcal{B}^\nu N_{\mu\nu} = \mathbf{B}^\dagger \mathbf{Y} - \frac{1}{2}\mathbf{B}^\dagger \mathbf{N}\mathbf{B} \,. \tag{56}$$

Here $\mathbf{Y}$ refers to the vector formed by the $y_\mu := \sum_I (x^I|\mathfrak{h}_\mu^I) = (\mathbf{x}|\mathfrak{h}_\mu)$, and $N_{\mu\nu} := \sum_I (\mathfrak{h}_\mu^I|\mathfrak{h}_\nu^I)$. We may write out y_μ explicitly in terms of x_μ:

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} x_1 - ix_2 + ix_3 + x_4 \\ x_1 + ix_2 + ix_3 - x_4 \\ x_1 - ix_2 - ix_3 - x_4 \\ x_1 + ix_2 - ix_3 + x_4 \end{bmatrix} \tag{57}$$

From the above it is clear that $y_4 = y_1^*, \ y_3 = y_2^*$.

Now we come to the quadratic term in the amplitudes. From the definitions of the basis functions given in Eq. (49), it is easy to directly compute the matrix $N_{\mu\nu}$. For example, for the CBC case we will have,

$$N_{11} = \sum_{I=1}^{M} (\mathfrak{h}_1^I|\mathfrak{h}_1^I)_I = \frac{1}{4}\sum_{I=1}^{M} (\mathfrak{f}^I \mathfrak{h}_0^{I\star}|\mathfrak{f}^I \mathfrak{h}_0^{I\star})_I = \frac{1}{2}\sum_{I=1}^{M} \sigma_I^2 |\mathfrak{f}^I|^2 \,, \tag{58}$$

where $\|\mathfrak{h}_0^I\|_I^2 := 2\sigma_I^2$ and $|\mathfrak{f}^I|^2$ is the magnitude of the complex number $\mathfrak{f}^I$. Similarly, it is easy to see that,

$$N_{12} = N_{21}^\star = \sum_{I=1}^{M} (\mathfrak{h}_1^I|\mathfrak{h}_2^I)_I = \frac{1}{4}\sum_{I=1}^{M} (\mathfrak{f}^I\mathfrak{h}_0^\star|\mathfrak{f}^{I\star}\mathfrak{h}_0^\star)_I = \frac{1}{2}\sum_{I=1}^{M} \sigma_I^2 (\mathfrak{f}^{I\star})^2 \,. \tag{59}$$

Note that here we have the square of $\mathfrak{f}^{I\star}$ rather than its norm. It is also easy to see for example that $N_{13} = N_{14} = 0$ and in fact, $\mathbf{N}$ has the usual block diagonal form:

$$\mathbf{N} = \frac{1}{2}\begin{bmatrix} \zeta & \kappa^\star & 0 & 0 \\ \kappa & \zeta & 0 & 0 \\ 0 & 0 & \zeta & \kappa^\star \\ 0 & 0 & \kappa & \zeta \end{bmatrix} , \tag{60}$$

Obtaining this block diagonal form is a useful, and perhaps surprising, benefit of using the complex amplitudes $\mathcal{B}^\mu$ rather than the real amplitudes of [9, 10] where the corresponding matrix is more complicated. The quantities ζ and κ appearing here are:

$$\zeta = \sum_{I=1}^{M} \sigma_I^2 |\mathfrak{f}^I|^2 \,, \qquad \kappa = \sum_{I=1}^{M} \sigma_I^2 (\mathfrak{f}^I)^2 \,. \tag{61}$$

One observes that ζ is real, while κ is complex in general. It is illustrative to re-express this in terms of the ψ-independent beam pattern functions a^I, b^I defined in Eq. (19):

$$\zeta = \sum_{I=1}^{M} \sigma_I^2 \left((a^I)^2 + (b^I)^2\right) \,, \qquad \kappa = \sum_{I=1}^{M} \sigma_I^2 (a^I + ib^I)^2 \,. \tag{62}$$

For detectors with identical PSDs, the σ_I will be identical as well. It is easy to see that ζ and κ are given by:

$$\zeta = |\mathbf{a}|^2 + |\mathbf{b}|^2 \,, \qquad \kappa = (\mathbf{a} + i\mathbf{b})^2 = |\mathbf{a}|^2 - |\mathbf{b}|^2 + 2i\mathbf{a}\cdot\mathbf{b} \,. \tag{63}$$

Here $\mathbf{a}$ is the vector $(a^1, a^2, \cdots, a^M)$ (and similarly for $\mathbf{b}$). The inner product is, as before, defined with the σ_I^2 as weights. Thus, for example, $\mathbf{a}\cdot\mathbf{b} = \sum_I \sigma_I^2 a^I b^I$ and $|\mathbf{a}|^2 := \mathbf{a}\cdot\mathbf{a}$.

5. The $\mathcal{B}$ statistic and its evaluation in the high SNR limit

We switch to the two complex variables $\mathcal{B}_1$ and $\mathcal{B}_2$ and write the log-likelihood function as:

$$\begin{aligned} \ln\Lambda = {} & \mathcal{B}_1^* y_1 + \mathcal{B}_1 y_1^* + \mathcal{B}_2 y_2 + \mathcal{B}_2^* y_2^* - \frac{1}{2}\zeta(|\mathcal{B}_1|^2 + |\mathcal{B}_2|^2) \\ & - \frac{1}{2}\mathcal{B}_1\mathcal{B}_2^*\kappa - \frac{1}{2}\mathcal{B}_1^*\mathcal{B}_2\kappa^* \equiv Q(\mathbf{Y};\mathbf{B}) \,. \end{aligned} \tag{64}$$

For the case $\kappa = 0$, the log likelihood can be written as sum of two terms as was shown in [9, 10]. In this case we can write $\ln\Lambda = \ln\Lambda_1 + \ln\Lambda_2$, where,

$$\ln\Lambda_1 = \mathcal{B}_1 y_1^* + \mathcal{B}_1^* y_1 - \frac{1}{2}\zeta|\mathcal{B}_1|^2, \quad \ln\Lambda_2 = \mathcal{B}_2 y_2^* + \mathcal{B}_2^* y_2 - \frac{1}{2}\zeta|\mathcal{B}_2|^2 \,. \tag{65}$$

For a uniform prior in the physical coordinates A, χ, φ_0, ψ, the Jacobian factor is as given in Eq. (53) and therefore, the $\mathcal{B}$ statistic is given by the integral:

$$\mathcal{B}(\mathbf{Y}) = \frac{1}{32} \int d\mathcal{B}_1 \; d\mathcal{B}_1^* \; d\mathcal{B}_2 \; d\mathcal{B}_2^* \; \left[\frac{e^{Q(\mathbf{Y};\mathbf{B})}}{(|\mathcal{B}_1|^2 |\mathcal{B}_2|^2)^{3/4}} \right] , \tag{66}$$

where $Q(\mathbf{Y};\mathbf{B})$ is given from Eq. (64). The integration range is determined from the range of A - this will be better seen when we write the same integral in polar coordinates - the limit then is on $|\mathcal{B}_1|$ and $|\mathcal{B}_2|$. When $\kappa = 0$, the above integral for this prior, because this prior factorizes similarly, also can be written as a product of two integrals, first one in $\mathcal{B}_1, \mathcal{B}_1^*$ and the second one in $\mathcal{B}_2, \mathcal{B}_2^*$. This agrees with the result in [9].

5.1. The $\mathcal{B}$ statistic integral in polar coordinates

We now go over to polar coordinates in which the integral appears more tractable. We write:

$$\mathcal{B}_1 = R_1 e^{i\theta_1}, \; \mathcal{B}_2 = R_2 e^{i\theta_2}, \; y_1 = \rho_1 e^{i\phi_1}, \; y_2 = \rho_2 e^{i\phi_2}, \; \kappa = k e^{4i\eta} . \tag{67}$$

Then the log likelihood can be written as follows:

$$\begin{aligned} \ln \Lambda = & -\frac{1}{2}\zeta(R_1^2 + R_2^2) + 2R_1\rho_1 \cos(\theta_1 - \phi_1) + 2R_2\rho_2 \cos(\theta_2 - \phi_2) \\ & - kR_1R_2 \cos(\theta_1 - \theta_2 + 4\eta) . \end{aligned} \tag{68}$$

The area element $d\mathcal{B}_1 d\mathcal{B}_1^* = -2iR_1 dR_1 d\theta_1$ and similarly for $d\mathcal{B}_2 d\mathcal{B}_2^*$. The integral now acquires a minus sign which can be corrected if we take the order of the physical variables as (A, ϕ, χ, ψ) instead of (A, χ, ϕ, ψ) (Or we can just take the modulus, if we do not care about the sign). With this change, the $\mathcal{B}$ statistic is:

$$\begin{aligned} & \mathcal{B}(\rho_1, \phi_1, \rho_2, \phi_2) = \frac{1}{8} \int dR_1 dR_2 d\theta_1 d\theta_2 \times \\ & \frac{e^{-\frac{1}{2}\zeta(R_1^2+R_2^2)+2R_1\rho_1\cos(\theta_1-\phi_1)+2R_2\rho_2\cos(\theta_2-\phi_2)-kR_1R_2\cos(\theta_1-\theta_2+4\eta)}}{\sqrt{R_1R_2}} . \end{aligned} \tag{69}$$

The range over which the integral is to be carried out can be worked out from the range allowed for the variables (A, χ, ϕ, ψ). Clearly, since ϕ, ψ cover the full range, so do the angles θ_1, θ_2 and so we have $0 \leq \theta_1, \theta_2 < 2\pi$. The range of R_1, R_2 is decided by the range of the amplitude A. If A is restricted to some range $A_{\rm min} < A < A_{\rm max}$, then only $A_{\rm max}$ matters and then we have $0 < R_1, R_2 < A_{\rm max}$. If there is no upper limit on A, then the integration region is a Cartesian product of two complex planes. Note that the integral is not singular, although the integrand is singular at the origin, because the singularity is weak.

5.2. The special case of $\kappa = 0$

When $\kappa = 0$, the integral in Eq. (69) can be written as a product of two integrals, one in ρ_1, ϕ_1 and the other in ρ_2, ϕ_2. Thus $\mathcal{B} = \text{const.} \times I_1 I_2$, where each of the integrals are of the form:

$$I = \int_0^\infty dR \;\; \frac{e^{-\frac{1}{2}\zeta R^2}}{\sqrt{R}} \int_0^{2\pi} d\theta \; e^{2\rho R \cos\theta} , \tag{70}$$

where ρ and ζ are constants. The integral over θ is the modified Bessel function I_0 and thus the integral can be written as,

$$I = 2\pi \int_0^\infty dR \, \frac{e^{-\frac{1}{2}\zeta R^2} I_0(2\rho R)}{\sqrt{R}} \, . \tag{71}$$

This agrees with the expression in [9] of Eq. (6.9). As shown therein, the integral can be written in terms of a confluent hypergeometric function.

In the limit of large ρR as will be the case if the threshold is high, the Bessel function can be replaced by its asymptotic form, which is valid when the argument of the Bessel function is greater than say, even 5. In the limit $x >> 1$, we have $I_0(x) \sim e^x/\sqrt{2\pi x}$, we use this asymptotic form to obtain:

$$\begin{aligned} I &\simeq \sqrt{\pi/\rho} \int_{R_0}^\infty dR \, \frac{e^{-\frac{1}{2}\zeta R^2 + 2\rho R}}{R} \, , \\ &= \sqrt{\pi/\rho} \; e^{\frac{2\rho^2}{\zeta}} \int_{R_0}^\infty dR \, \frac{e^{-\frac{1}{2}\zeta (R - 2\rho/\zeta)^2}}{R} \, . \end{aligned} \tag{72}$$

where the asymptotic form of the I_0 is valid for $R > R_0$. Also it is assumed that the contribution to the integral is negligible for $R < R_0$. Further, the R in the denominator can be replaced by its mean value in the Gaussian, namely, $2\rho/\zeta$ and then the Gaussian integral performed assuming $2\rho/\zeta - R_0 >> (\zeta)^{-1/2}$ to yield,

$$I \sim \pi \left(\frac{\zeta}{2} \right)^{1/2} \rho^{-3/2} \, e^{\frac{2\rho^2}{\zeta}} \, . \tag{73}$$

Collecting all the terms and including the constant, the $\mathcal{B}$ statistic is given by,

$$\mathcal{B}(\rho_1, \rho_2) \simeq \frac{\pi^2}{16} \, \zeta \, [\rho_1 \rho_2]^{-3/2} \, e^{\frac{2}{\zeta}(\rho_1^2 + \rho_2^2)} \, . \tag{74}$$

Recall that $\rho_1 = |y_1|$ and $\rho_2 = |y_2|$. This must agree with the results obtained in [9] if one takes the asymptotic form of the confluent hypergeometric function as obtained therein.

5.3. The general case of $\kappa \neq 0$

Here the basic technique involves resorting to the principle axes transformation which diagonalizes the matrix $\mathbf{N}$ given in Eq. (60) and makes the quadratic form $\mathbf{B}^\dagger \mathbf{N} \mathbf{B}$ a sum of squares. We observe here that $\mathbf{N}$ is block diagonal with the *same* block repeated. Thus we need to just focus on one block which is just a 2×2 matrix - we call this matrix $\mathbf{N}_2$:

$$\mathbf{N}_2 = \begin{bmatrix} \zeta & \kappa^* \\ \kappa & \zeta \end{bmatrix} , \tag{75}$$

Recalling $\kappa = k e^{4i\eta}$, the eigenvalues of $\mathbf{N}_2$ are given by $\zeta \pm k$. The matrix $\mathbf{U}$ diagonalizing $\mathbf{N}_2$ is unitary and is given by:

$$\mathbf{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} e^{-2i\eta} & -e^{-2i\eta} \\ e^{2i\eta} & e^{2i\eta} \end{bmatrix} \quad \text{where} \quad \tan 4\eta = \frac{2(\mathbf{a} \cdot \mathbf{b})}{|\mathbf{a}|^2 - |\mathbf{b}|^2} \, . \tag{76}$$

From the above, we have the result:

$$\mathbf{U}^\dagger \mathbf{N}_2 \mathbf{U} = \begin{bmatrix} \zeta + k & 0 \\ 0 & \zeta - k \end{bmatrix} . \tag{77}$$

The diagonalisation process leads us to another set of amplitude variables C_1, C_2, where then,

$$\begin{bmatrix} \mathcal{B}_1 \\ \mathcal{B}_2 \end{bmatrix} = \mathbf{U} \begin{bmatrix} C_1 \\ C_2 \end{bmatrix} , \tag{78}$$

which gives,

$$\mathcal{B}_1 = \frac{1}{\sqrt{2}} e^{-2i\eta}(C_1 - C_2) , \qquad \mathcal{B}_2 = \frac{1}{\sqrt{2}} e^{2i\eta}(C_1 + C_2) . \tag{79}$$

The inverse transformation is

$$C_1 = \frac{1}{\sqrt{2}} \left(e^{2i\eta}\mathcal{B}_1 + e^{-2i\eta}\mathcal{B}_2 \right) , \qquad C_2 = \frac{1}{\sqrt{2}} \left(-e^{2i\eta}\mathcal{B}_1 + e^{-2i\eta}\mathcal{B}_2 \right) . \tag{80}$$

Given the expressions of Eq. (48) for the $\mathcal{B}_\mu$, we can view this as a (sky-dependent) shift in the polarization angle $\psi \to \psi - \eta(\theta, \phi)$. This is a re-derivation of the so-called dominant polarization frame discussed in Eq. (2.33) of [5].

Similarly, we define the complex data vectors z_1, z_2 by,

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \mathbf{U}^\dagger \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} . \tag{81}$$

Therefore, $z_1 = \alpha y_1 + \alpha^* y_2$ and $z_2 = -\alpha y_1 + \alpha^* y_2$.

Now we can write the log likelihood in terms of the new amplitudes C_μ and z_μ. We then have:

$$\ln \Lambda = C_1^* z_1 + C_2^* z_2 + C_1 z_1^* + C_2 z_2^* - \frac{1}{2}(\zeta + k)|C_1|^2 - \frac{1}{2}(\zeta - k)|C_2|^2 . \tag{82}$$

Now no cross terms in C_1, C_2 appear and the expression for the log likelihood is much simpler. However, the Jacobian factor in the denominator gets more complex because

$$|B_1|^2 |B_2|^2 = \frac{1}{4} |C_1 - C_2|^2 |C_1 + C_2|^2 . \tag{83}$$

The Jacobian:

$$\frac{\partial(B_1, B_1^*, B_2, B_2^*)}{\partial(C_1, C_1^*, C_2, C_2^*)} = 1 , \tag{84}$$

and hence the volume elements are $dB_1 dB_1^* dB_2 dB_2^* = dC_1 dC_1^* dC_2 dC_2^*$ since $\mathbf{U}$ is unitary. We now have all the ingredients to write down the $\mathcal{B}$ statistic:

$$\mathcal{B}(z_\mu) = \frac{\sqrt{2}}{16} \int \frac{dC_1 dC_1^* dC_2 dC_2^*}{|C_1 - C_2|^{3/2} |C_1 + C_2|^{3/2}} \times e^{C_1^* z_1 + C_2^* z_2 + C_1 z_1^* + C_2 z_2^* - \frac{1}{2}(\zeta + k)|C_1|^2 - \frac{1}{2}(\zeta - k)|C_2|^2} . \tag{85}$$

We first consider the case when the signal-to-noise-ratio (SNR) is sufficiently high so that we can focus on the peak of the terms in the exponent. Collecting terms in C_1 and C_2, writing $\ln\Lambda = \ln\Lambda_+ + \ln\Lambda_-$, and completing squares in $\ln\Lambda_\pm$:

$$\ln\Lambda_\pm = \frac{2|z|^2}{\zeta \pm k} - \frac{1}{2}(\zeta \pm k)\left|C - \frac{2z}{\zeta \pm k}\right|^2 . \tag{86}$$

Here it is assumed that in C and z are respectively C_1 and z_1 in $\ln\Lambda_+$, and C_2 and z_2 in $\ln\Lambda_2$. Under the simplest approximation, in the denominator of the integrand in Eq. (85), we set $C_{1,2}$ to the values at the peak of the Gaussian

$$C_1 = \frac{2z_1}{\zeta + k}, \quad C_2 = \frac{2z_2}{\zeta - k} . \tag{87}$$

Performing the two of Gaussian integrals yields (taking the integrals to be over the whole complex plane)

$$\mathcal{B} = \frac{4\pi^2}{(\zeta^2 - k^2)}\left|\frac{4z_1^2}{(\zeta + k)^2} - \frac{4z_2^2}{(\zeta - k)^2}\right|^{-3/2} \exp\left[\frac{2|z_1|^2}{(\zeta + k)} + \frac{2|z_2|^2}{(\zeta - k)}\right] \tag{88}$$

We first check that this expression reduces to the special case of $\kappa = 0$ derived in section 5.2, Eq. (74). In this case, we put $k = 0$ in Eq. (88). We first note that z_1, z_2 are related to y_1, y_2 by a unitary transformation $\mathbf{U}$ given in Eq. (81). Taking the Hermitian conjugate of this equation and multiplying this with the equation itself, it follows that $|z_1|^2 + |z_2|^2 = |y_1|^2 + |y_2|^2$. This takes care of the exponential term. For the denominator, we see that $|z_1 - z_2| = \sqrt{2}|y_1|$ and $|z_1 + z_2| = \sqrt{2}|y_2|$. Again collecting all terms we finally deduce that:

$$\mathcal{B} = \frac{\pi^2}{16}\,\zeta\,\frac{e^{\frac{2}{\zeta}\left(|y_1|^2 + |y_2|^2\right)}}{(|y_1||y_2|)^{3/2}} , \tag{89}$$

which is in exact agreement with Eq. (74).

Also from the expression of κ, we find $k^2 = |\kappa|^2 = |\mathbf{a}|^4 + |\mathbf{b}|^4 + 2|\mathbf{a}|^2|\mathbf{b}|^2\cos 2\theta \leq (|\mathbf{a}|^2 + |\mathbf{b}|^2)^2 \equiv \zeta^2$, where we have written $\mathbf{a}\cdot\mathbf{b} = |\mathbf{a}||\mathbf{b}|\cos\theta$. This means that $k \leq \zeta$. The $\mathcal{B}$ statistic is singular when $\zeta = k$. In fact, setting $\epsilon = \zeta - k$ with ϵ small, we see that

$$\mathcal{B} \sim \epsilon^2 e^{2|z_2|^2/\epsilon} . \tag{90}$$

This is singular when $\epsilon \to 0$.

5.4. Singularities of the $\mathcal{B}$ statistic

The singularities of the $\mathcal{B}$ statistic given by Eq. (88) occur when $\zeta = k$. We can see from Eq. (88) that it will *also* be singular when:

$$\frac{z_1}{\zeta + \kappa} = \pm\frac{z_2}{\zeta - \kappa} . \tag{91}$$

These correspond to the two cases where the maximum-likelihood signal is purely circularly polarized, either left or right. Whereas the singularities that occur when

$\zeta = \pm k$ are characteristic of the detector network alone, the singularities above depend also on the data, and the extent to which it prefers pure circular polarization.

Let us now consider the first kind of singularity. To avoid clutter, we drop the factors of σ_I which account for the detector sensitivities (or alternatively, scale the beam pattern functions appropriately – this will not affect the present discussion). Given that $\zeta \geq k$ and $\zeta, k \geq 0$, we see that $\zeta + k = 0$ if and only if $\zeta = 0$. Consider then the quantity $\zeta^2 - k^2$. If this quantity vanishes then either $\zeta = 0$ or $\zeta = k$. Thus, away from sky-positions where $\zeta = 0$, which is equivalent to both a and b vanishing for each detector, we see that the set of points where $\zeta^2 - k^2$ vanishes contains all such singular points. A straightforward calculation yields

$$\begin{aligned}\zeta^2 - k^2 &= \left(\sum_I |\mathfrak{f}_I|^2\right)^2 - \left|\sum_I \mathfrak{f}_I^2\right|^2 = \sum_{I,J}\left(\mathfrak{f}_I\mathfrak{f}_I^\star\mathfrak{f}_J\mathfrak{f}_J^\star - \mathfrak{f}_I^2\mathfrak{f}_J^{\star 2}\right)\\ &= -\sum_{I>J}\left(\mathfrak{f}_I\mathfrak{f}_J^\star - \mathfrak{f}_I^\star\mathfrak{f}_J\right)^2 .\end{aligned} \tag{92}$$

The terms with $I = J$ vanish and we note also that the quantity $\mathfrak{f}_I\mathfrak{f}_J^\star - \mathfrak{f}_I^\star\mathfrak{f}_J$ is pure imaginary. Thus its square is negative and we get

$$\zeta^2 - k^2 = \sum_{I>J}\left|\mathfrak{f}_I\mathfrak{f}_J^\star - \mathfrak{f}_I^\star\mathfrak{f}_J\right|^2 . \tag{93}$$

We have thus written $\zeta^2 - k^2$ as a sum of positive terms and as a sum over detector pairs. If $\zeta^2 - k^2 = 0$, then each term in the above sum must vanish. Consider now a single term from this expression:

$$|\mathfrak{f}_I\mathfrak{f}_J^\star - \mathfrak{f}_I^\star\mathfrak{f}_J| = |\mathfrak{f}_I\mathfrak{f}_J| \times \left|\frac{\mathfrak{f}_J^\star}{\mathfrak{f}_J} - \frac{\mathfrak{f}_I^\star}{\mathfrak{f}_I}\right| . \tag{94}$$

We see immediately that apart from points where $\mathfrak{f}_I$ and $\mathfrak{f}_J$ vanish, the degeneracy occurs when $\frac{\mathfrak{f}_J^\star}{\mathfrak{f}_J} = \frac{\mathfrak{f}_I^\star}{\mathfrak{f}_I}$. This is just the condition that the arguments of $\mathfrak{f}_I$ and $\mathfrak{f}_J$ are equal (or differ by π). The condition for $\zeta^2 - k^2$ to vanish is then that the arguments for *all* $\mathfrak{f}_I$ should be equal (modulo π). For two detectors, we will thus generically get curves on the celestial sphere. For three detectors, we will get a finite number of degenerate points (the intersection of the curves), and no degeneracy generically for 4 or more detectors.

The second kind of singularity can also be discussed in terms of the quantity $\zeta^2 - k^2$ but is more complicated due to the presence of the data. We shall leave further discussion of the geometric nature of the singularity to future work and turn now to the properties of the $\mathcal{B}$ statistic at the singularity.

5.5. Applying the Laplace approximation to the $\mathcal{B}$ statistic

We can formalize the approximations to the integral that we have been making through use of *Laplace's approximation.* As the name suggests, the technique is an old one. In its most familiar form, it states that if a function f takes on a unique, non-degenerate minimum at the origin within some domain R of $\mathbb{R}^n$ that properly contains the origin, and that if the value of that minimum is zero, then:

$$\int_R e^{-\lambda f}\, g(x)\, d^n\mathbf{x} \sim \left(\frac{2\pi}{\lambda}\right)^{n/2} \frac{g(\mathbf{0})}{\sqrt{\det Hf(\mathbf{0})}} \tag{95}$$

for sufficiently large λ. Here $Hf(\mathbf{0})$ denotes the Hessian of f, evaluated at the origin. This approximation, and the related stationary-phase approximation, are well known in applied mathematics and we have used them already in this paper. They encapsulate the idea that for a sharply peaked integrand, the value of the integral should be chiefly determined by its behavior at its peak.

Much less well known is that the approximation of Eq. (95) is just the first term in a full asymptotic expansion. Moreover, the requirements that the minimum be non-degenerate, and the evident requirement that g be well-defined at the origin, may both be relaxed. All that is really required is that both f and g themselves possess asymptotic expansions, as $|\mathbf{x}| \to \mathbf{0}$, and that any singularities of g should be integrable. There is extensive discussion of these topics in the book by Wong [20], but we shall refer primarily to the paper [21] by Kirwin, which collects in one place the specific generalizations of Laplace's method that we shall need.

We will focus only on the leading order term in the asymptotic expansion; our interest in generalizing is to relax the assumption that the prior (that part of the integrand not in the exponential) be finite: we will instead allow it an integrable singularity. As we shall see, this is necessary to apply Laplace's approximation to the case where the gravitational wave strain data is peaked at a pure circular polarization state.

So, we still assume that the function f in Eq. (95) has a unique minimum value at the origin, and that the value of that minimum is zero. We do *not* assume that the minimum is non-degenerate. If $\mathbf{x} = (x^1, \ldots, x^n)$ are Cartesian coordinates on $\mathbb{R}^n$, then introduce generalized spherical coordinates by defining $r = \sqrt{(x^1)^2 + \cdots (x^n)^2}$, and denoting by S^{n-1} the unit sphere $r = 1$; also define $\Omega = \mathbf{x}/r$. The hypotheses that are required to apply the results of [21] are:

(i) The function f possesses $N+1$ continuous functions f_j on Ω, with $f_0(\Omega) > 0$, such that for some real number $\nu > 0$:

$$f(r,\Omega) = r^{\nu} \sum_{j=0}^{N} f_j(\Omega) r^j + o(r^{N+\nu}) \qquad \text{as } r \to 0, \text{ and} \tag{96}$$

(ii) The function g possesses $N+1$ continuous functions g_j on Ω, such that for some real number $\mu > 0$:

$$g(r,\Omega) = r^{\mu-n} \sum_{j=0}^{N} g_j(\Omega) r^j + o(r^{N+\mu-n}) \qquad \text{as } r \to 0, \text{ and} \tag{97}$$

(iii) There exists some $\lambda_0 > 0$ such that the integral

$$\int_R e^{-\lambda_0 f}\, g\, d^n x \tag{98}$$

converges.

When these conditions hold, the leading term of the asymptotic expansion is given by:

$$\int_R e^{-\lambda f}\, g\, d^n x \;\sim\; \lambda^{-\mu/\nu}\frac{1}{\nu}\Gamma\left(\frac{\mu}{\nu}\right)\int_{S^{n-1}}\frac{g_0(\Omega)}{[f_0(\Omega)]^{\mu/\nu}}\,d\Omega \tag{99}$$

Though it is not obvious from Eq. (99), in the familiar case where the minimum of f is non-generate (implying the determinant of the Hessian of f at its minimum is nonzero) and when $\mu = n$, so that g has no singularity at the minimum of f, then Eq. (99) reduces to Eq. (95).

To apply either Eq. (95) or Eq. (99), we may use either integral form of the $\mathcal{B}$ statistic, whether Eq. (66) or Eq. (85). In either case, we must rewrite the expression so that the minimum of our log-likelihood occurs at the origin, and the value of that minimum must be zero. This is accomplished in each case by completing the square.

If we start with Eq. (66), define:

$$\hat{\mathbf{B}} = \mathbf{N}^{-1}\mathbf{Y}, \qquad \Delta\mathbf{B} = \mathbf{B} - \hat{\mathbf{B}} \tag{100}$$

Then it is easy to show that:

$$Q(\mathbf{Y};\mathbf{B}) = -\frac{1}{2}\Delta\mathbf{B}^\dagger\mathbf{N}\Delta\mathbf{B} + \frac{1}{2}\hat{\mathbf{B}}^\dagger\mathbf{N}\hat{\mathbf{B}} \tag{101}$$

Using this, we get:

$$\mathcal{B}(\mathbf{Y}) = \frac{e^{\frac{1}{2}\hat{\mathbf{B}}^\dagger\mathbf{N}\hat{\mathbf{B}}}}{32}\int\frac{d(\Delta\mathcal{B}_1)\,d(\Delta\mathcal{B}_1^\star)\,d(\Delta\mathcal{B}_2)\,d(\Delta\mathcal{B}_2^\star)\,e^{-\frac{1}{2}\Delta\mathbf{B}^\dagger\mathbf{N}\Delta\mathbf{B}}}{\left[(\Delta\mathcal{B}_1+\hat{\mathcal{B}}_1)(\Delta\mathcal{B}_1^\star+\hat{\mathcal{B}}_1^\star)(\Delta\mathcal{B}_2+\hat{\mathcal{B}}_2)(\Delta\mathcal{B}_2^\star+\hat{\mathcal{B}}_2^\star)\right]^{3/4}} \tag{102}$$

We now have an exponential in the integrand which has its minimum value at the origin, and a minimum value of zero. It is also helpful to explicitly include a parameter that will be large when, as we consider, there is a threshold for triggers. Define the following:

$$\begin{aligned}\lambda &= \sqrt{|\hat{\mathcal{B}}_1|^2 + |\hat{\mathcal{B}}_2|^2},\\ \mathbf{U} &= \Delta\mathbf{B}/\lambda,\\ \hat{\mathbf{U}} &= \hat{\mathbf{B}}/\lambda.\end{aligned}$$

Using this, we can write:

$$\mathcal{B}(\mathbf{Y}) = \lambda\frac{e^{\frac{1}{2}\hat{\mathbf{B}}^\dagger\mathbf{N}\hat{\mathbf{B}}}}{32}\int\frac{d(U_1)\,d(U_1^\star)\,d(U_2)\,d(U_2^\star)\,e^{-\frac{\lambda^2}{2}\mathbf{U}^\dagger\mathbf{N}\mathbf{U}}}{\left[(U_1+\hat{U}_1)(U_1^\star+\hat{U}_1^\star)(U_2+\hat{U}_2)(U_2^\star+\hat{U}_2^\star)\right]^{3/4}} \tag{103}$$

Note that $|\hat{U}_1|^2 + |\hat{U}_2|^2 = 1$.

It is now straightforward to consider the cases of Laplace's transformation that we may apply to this expression. If neither of $|\hat{U}_1|$ or $|\hat{U}_2|$ are zero, then the function g in our formulation of Laplace's approximation is not singular at the origin, and we may directly apply Eq. (95) to get:

$$\mathcal{B}(\mathbf{Y}) \sim \frac{\pi^2}{2(\zeta^2 - k^2)}\frac{e^{\frac{1}{2}\hat{\mathbf{B}}^\dagger\mathbf{N}\hat{\mathbf{B}}}}{\left(|\hat{\mathcal{B}}_1||\hat{\mathcal{B}}_2|\right)^{3/2}}. \tag{104}$$

and where we remind the reader that the functional dependence of the right-hand side on $\mathbf{Y}$ is through $\hat{\mathbf{B}} = \mathbf{N}^{-1}\mathbf{Y}$.

However, if the peak value of one of $|\hat{U}_1|$ or $|\hat{U}_2|$ does vanish, then we must use Eq. (99); this happens if the signal is purely circularly polarized. For concreteness, suppose that $\hat{U}_1 = 0$; note that this then implies $|\hat{U}_2| = 1$. In that case, as $r \to 0$, the factors in g that contain $\hat{U}_1$ diverge, though the singularity is integrable. We have:

$$\begin{aligned} f(r,\Omega) &= \mathbf{U}^\dagger \mathbf{N} \mathbf{U} \\ &= r^2 \left[\zeta + 2k \sin\beta\, \cos\beta \cos(\alpha_1 - \alpha_2 + 4\eta)\right] \end{aligned} \tag{105}$$

In this equation, we have introduced Hopf coordinates on S^3:

$$U_1 = r\cos\beta\, e^{i\alpha_1}\,, \quad U_2 = r\sin\beta\, e^{i\alpha_2}\,, \quad d\Omega = \sin\beta\cos\beta\, d\beta\, d\alpha_1 d\alpha_2. \tag{106}$$

In these same coordinates, we have:

$$g(r,\Omega) \to \left[(r\cos\beta\, e^{i\alpha_1})(r\cos\beta\, e^{-i\alpha_1})\right]^{-3/4} = r^{-3/2}(\cos\beta)^{-3/2} \tag{107}$$

From this we conclude that the exponent μ is $5/2$; it is also immediate that the exponent ν is two, since our function f is a quadratic form. Thus, if we put together all of the elements of Eq. (99), we obtain in this case:

$$\mathcal{B}(\mathbf{Y}) \sim \frac{2^{1/4}}{32\zeta^{5/4}|\hat{\mathcal{B}}_2|^{3/2}} \Gamma\left(\frac{5}{4}\right) e^{\frac{1}{2}\hat{\mathbf{B}}^\dagger \mathbf{N} \hat{\mathbf{B}}} I\left(\frac{k}{\zeta}\right) \tag{108}$$

where we have defined the function $I(x)$, for $0 \leq x < 1$, by the three-dimensional integral:

$$\begin{aligned} I(x) = &\int_0^{\pi/2} d\beta \int_0^{2\pi} d\alpha_1 \int_0^{2\pi} d\alpha_2 \\ &\times \frac{\sin\beta\cos\beta}{[\cos\beta]^{3/2}\left[1 + 2x\sin\beta\cos\beta\cos\left(\alpha_1 - \alpha_2 + 4\eta\right)\right]^{5/4}} \end{aligned} \tag{109}$$

Though this integral may appear intractable, in fact it may be performed exactly, as outlined in Appendix B. The result is:

$$I(x) = 8\pi^2(1 - x^2)^{-1/4} \tag{110}$$

and therefore we obtain:

$$\mathcal{B}(\mathbf{Y}) \sim \frac{2^{1/4}\Gamma\left(\frac{1}{4}\right)\pi^2}{16\zeta^{3/4}(\zeta^2 - k^2)^{1/4}} \frac{e^{\frac{1}{2}\hat{\mathbf{B}}^\dagger \mathbf{N} \hat{\mathbf{B}}}}{|\hat{\mathcal{B}}_2|^{3/2}} \tag{111}$$

If we compare this equation to Eq. (104), we see that the exponential factor in Eq. (111) is identical to that in Eq. (111), as is the dependence on $\hat{\mathcal{B}}_2$. The only functional difference in the dependence on the data is that the singularity that Eq. (104) would predict as $\hat{\mathcal{B}}_1 \to 0$ is absent from Eq. (111).

It would seem desirable to have an expression that smoothly interpolates between Eqs. (104) and (111). This is a considerably more complex problem. The essential difficulty is that we are attempting an asymptotic approximation not of a single integral, but rather a family of integrals, depending on parameters (in our case, the $\mathbf{Y}$ and through those the $\hat{\mathbf{B}}$). To smoothly interpolate, we would need an asymptotic approximation that was *uniform* in these parameters. This topic is discussed in depth in chapter seven of Wong's book [20] and references cited therein, as well as a different approach in [22], but we have found no simple application of those techniques to our problem, and defer an investigation of that for future work.

6. Discussion

In this paper, we have studied the analytic marginalization of the likelihood function over the nuisance parameters $(A, \varphi_0, \iota, \psi)$ for modeled gravitational wave searched. The results are applicable for both transient binary coalescence signals, and the long duration signals emitted by rapidly rotating neutron stars. The main theme of this paper is that the matrix elements of the rotation group T^{ℓ}_{mn} are useful amplitude parameters. Not only are they natural from a geometric viewpoint, but they also simplify the calculation of the $\mathcal{B}$-statistic. The same interpretation is useful also in the calculation of the beam pattern function for a gravitational wave detector. Thus, it is natural to view the coordinates φ_0, ψ, ι as coordinates on the group of rotations or equivalently, on S^3. Using these coordinates, we have derived useful approximations to the $\mathcal{B}$-statistic in the case when the SNR is high. We have investigated the singularities of the detector network and we have obtained a simple expression for $\mathcal{B}$ in the singular case.

We have restricted ourselves to non-precessing systems and in fact we have also not considered higher modes in the waveform model. In both these case, we will need to go beyond the dominant mode of the gravitational wave signal. It is clear that these higher modes can be included in our formalism quite naturally. It also seems plausible that extensions of the methods used here might also work in those cases. It is an accident that for non-precessing systems, the number of amplitude parameters is the same as the number of physical parameters. Including precession or higher modes will result in a larger number of amplitude parameters than the number of physical parameters. The Bayesian framework naturally deals with this mismatch in the number of parameters. Even restricting ourselves to non-precessing systems and ignoring the higher modes, the results of this paper might be useful in parameter estimation methods and in suggesting modifications to the detection statistic for the searches as well.

Acknowledgments

We are grateful to Archana Pai for valuable comments. JLW acknowledges support from NSF award PHY-1506254.

References

[1] Stuart A, Ord K and Arnold S 2007 *Classical Inference and the Linear Model* Kendall's Advanced Theory of Statistics, Volume 2A, Sixth Edition (Hodder Headline Group) ISBN 9780340662304
[2] Jaranowski P, Krolak A and Schutz B F 1998 *Phys. Rev.* **D58** 063001 (*Preprint* `gr-qc/9804014`)

[3] Cutler C and Schutz B F 2005 *Phys. Rev.* **D72** 063006 (*Preprint* gr-qc/0504011)
[4] Pai A, Dhurandhar S and Bose S 2001 *Phys. Rev.* **D64** 042004 (*Preprint* gr-qc/0009078)
[5] Harry I W and Fairhurst S 2011 *Phys. Rev.* **D83** 084002 (*Preprint* 1012.4939)
[6] Keppel D 2013 (*Preprint* 1307.4158)
[7] Searle A C 2008 Monte-Carlo and Bayesian techniques in gravitational wave burst data analysis *Proceedings, 12th Workshop on Gravitational wave data analysis (GWDAW-12): Cambridge, USA, December 13-16, 2007* (*Preprint* 0804.1161) URL http://inspirehep.net/record/782996/files/arXiv:0804.1161.pdf
[8] Prix R and Krishnan B 2009 *Class. Quant. Grav.* **26** 204013 (*Preprint* 0907.2569)
[9] Whelan J T, Prix R, Cutler C J and Willis J L 2014 *Class. Quant. Grav.* **31** 065002 (*Preprint* 1311.0065)
[10] Haris K and Pai A 2016 (*Preprint* 1606.03867)
[11] Schutz B F and Tinto M 1987 *MNRAS* **224** 131–154
[12] Estabrook F B and Wahlquist H D 1975 *Gen. Rel. and Grav.* **6** 439–447
[13] Finn L S 2009 *Phys. Rev.* **D79** 022002 (*Preprint* 0810.4529)
[14] Dhurandhar S V and Tinto M 1988 *Month. Not. Roy. Astron. Soc.* **234** 663–676
[15] Goldberg J N, MacFarlane A J, Newman E T, Rohrlich F and Sudarshan E C G 1967 *J. Math. Phys.* **8** 2155
[16] Newman E and Penrose R 1962 *J. Math. Phys.* **3** 566–578
[17] Gelfand I M, Minlos R A and Shapiro Z Y 1963 *Representations of the rotation and Lotentz groups and their applications* (New York: Pergamon Press) ISBN 978-1614273462
[18] Baskaran D and Grishchuk L P 2004 *Class. Quant. Grav.* **21** 4041–4062 (*Preprint* gr-qc/0309058)
[19] Peters P C and Mathews J 1963 *Phys. Rev.* **131** 435–439
[20] Wong R 2001 *Asymptotic Approximations of Integrals* Classics in Applied Mathematics (Society for Industrial and Applied Mathematics) ISBN 9780898714975
[21] Kirwin W D 2010 *Asymp. Anal.* **70** 231–248
[22] Connor J N L 1973 *Faraday Discuss. Chem. Soc.* **55**(0) 51–58 URL http://dx.doi.org/10.1039/DC9735500051
[23] Meurer A, Smith C P, Paprocki M, Čertík O, Kirpichev S B, Rocklin M, Kumar A, Ivanov S, Moore J K, Singh S, Rathnayake T, Vig S, Granger B E, Muller R P, Bonazzi F, Gupta H, Vats S, Johansson F, Pedregosa F, Curry M J, Terrel A R, Roučka v, Saboo A, Fernando I, Kulal S, Cimrman R and Scopatz A 2017 *PeerJ Computer Science* **3** e103 ISSN 2376-5992 URL https://doi.org/10.7717/peerj-cs.103

Appendix A. The matrix elements of the $\ell = 2$ representation of the rotation group

A rotation g is represented by matrices $T^\ell_{mn}[g]$ where ℓ is the weight of the representation and m, n are the indices for the matrix elements with $-\ell \leq m, n \leq \ell$. If we use the Euler angles (α, β, γ) to parametrize a rotation then T^ℓ_{mn} are functions of (α, β, γ). Following [17] we shall use the 'zxz' convention for the Euler angles. Thus we shall go from a frame (x, y, z) to (x', y', z') starting with a rotation around the z axis by an angle α, a rotation around the x-axis by an angle β, and finally a rotation around the z axis by an angle γ. The T^ℓ_{mn} are also called the Wigner D-matrices, and the spin weighted spherical harmonics ${}_sY_{\ell m}(\theta, \phi)$ are proportional to $T^\ell_{-s,m}(\theta, \phi, 0)$, i.e. with the third Euler angle set to zero [15].

It can be shown that

$$T^\ell_{mn}(\alpha, \beta, \gamma) = e^{-im\alpha} P^\ell_{mn}(\cos\beta) e^{-in\gamma}\,, \tag{A.1}$$

where $P^{\ell}_{mn}(\cos\beta)$ are the Jacobi polynomials:

$$P^{\ell}_{mn}(x) = \frac{(-1)^{\ell-m} i^{n-m}}{2^{\ell}(\ell-m)!}\sqrt{\frac{(\ell-m)!(\ell+n)!}{(\ell+m)!(\ell-n)!}} \times (1-x)^{-\frac{n-m}{2}}(1+x)^{-\frac{n+m}{2}}\frac{d^{\ell-n}}{dx^{\ell-n}}[(1-x)^{\ell-m}(1+x)^{\ell+m}]\,. \tag{A.2}$$

It will be useful to list explicitly the $\ell = 2$ matrix elements for $m = -2$

$$P^{2}_{-2,-2}(\cos\beta) = \frac{1}{4}(1+\cos\beta)^2\,, \tag{A.3}$$

$$P^{2}_{-2,-1}(\cos\beta) = \frac{i}{2}\sin\beta(1+\cos\beta)\,, \tag{A.4}$$

$$P^{2}_{-2,0}(\cos\beta) = -\frac{1}{2}\sqrt{\frac{3}{2}}(1-\cos^2\beta)\,, \tag{A.5}$$

$$P^{2}_{-2,1}(\cos\beta) = \frac{i}{2}\sin\beta(\cos\beta-1)\,, \tag{A.6}$$

$$P^{2}_{-2,2}(\cos\beta) = \frac{1}{4}(1-\cos\beta)^2 \tag{A.7}$$

Note that

$$P^{\ell}_{mn}(x) = P^{\ell\,\star}_{nm}(x)\,. \tag{A.8}$$

Appendix B. Evaluating the integral $I(x)$

Here we outline the steps in the closed-form evaluation of the function $I(x)$ defined in Eq. (109), which we repeat here for convenience:

$$I(x) = \int_0^{\pi/2} d\beta \int_0^{2\pi} d\alpha_1 \int_0^{2\pi} d\alpha_2 \times \frac{\sin\beta\cos\beta}{[\cos\beta]^{3/2}\left[1+2x\sin\beta\cos\beta\cos\left(\alpha_1-\alpha_2+4\eta\right)\right]^{5/4}} \tag{B.1}$$

To begin our evaluation of this integral, first change variables from α_1 and α_2 to new variables γ_+ and γ_- defined by:

$$\gamma_+ = \alpha_1+\alpha_2, \qquad \gamma_- = \alpha_1-\alpha_2 \tag{B.2}$$

It is easy to check that $d\alpha_1 d\alpha_2 = \frac{1}{2}d\gamma_+ d\gamma_-$; we must also adjust the limits of integration. Our goal is to perform the integral over γ_+ trivially, so we write:

$$I(x) = \frac{1}{2}\int_0^{\pi/2} d\beta \int_{-2\pi}^{2\pi} d\gamma_- \int_{-2\pi+|\gamma_-|}^{2\pi-|\gamma_-|} d\gamma_+ \times \frac{\sin\beta\cos\beta}{[\cos\beta]^{3/2}\left[1+2x\sin\beta\cos\beta\cos\left(\gamma_-+4\eta\right)\right]^{5/4}} \tag{B.3}$$

$$= \frac{1}{2}\int_0^{\pi/2} d\beta \int_{-2\pi}^{2\pi} d\gamma_-(4\pi - 2|\gamma_-|) \times \frac{\sin\beta\cos\beta}{[\cos\beta]^{3/2}\left[1+2x\sin\beta\cos\beta\cos\left(\gamma_-\right)\right]^{5/4}} \tag{B.4}$$

In going from Eq. (B.4) to Eq. (B.5), we have performed the integral over γ_+. It is subtler to see that we may drop the shift by 4η, but it is possible to check that the potential extra term in the $|\gamma_-|$ in fact vanishes when integrated from -2π to 2π.

Next, we use the identity:

$$(1-z)^{-a} = \sum_{k=0}^{\infty} \frac{(a)_k}{k!} z^k, \qquad |z| < 1 \tag{B.5}$$

where $(a)_k = \Gamma(a+k)/\Gamma(a)$ is the Pochhammer symbol, to expand the binomial fractional power in the integrand; this expansion will be valid for $0 \leq x < 1$ as we require. This gives:

$$I(x) = \frac{1}{2} \sum_{k=0}^{\infty} \frac{(\frac{5}{4})_k}{k!} (-2x)^k \int_{-2\pi}^{2\pi} (2\pi - |\gamma_-|)(\cos\gamma_-)^k d\gamma_- \times \int_0^{\pi/2} \frac{(\sin\beta)^{k+1}(\cos\beta)^{k+1}}{[\cos\beta]^{3/2}} d\beta. \tag{B.6}$$

The integral over γ_- can now be performed by elementary methods, and vanishes when k is odd. When $k = 2m$ is even, it is:

$$\int_{-2\pi}^{2\pi} (2\pi - |\gamma_-|)(\cos\gamma_-)^{2m} d\gamma_- = \frac{4\pi^2}{2^{2m}} \binom{2m}{m} \tag{B.7}$$

and inserting this gives:

$$I(x) = 4\pi^2 \sum_{m=0}^{\infty} \frac{(\frac{5}{4})_{2m}}{(2m)!} \binom{2m}{m} x^{2m} \int_0^{\pi/2} (\sin\beta)^{2m+1} (\cos\beta)^{2m-\frac{1}{2}} d\beta \tag{B.8}$$

Now we define $u = \cos\beta$ to re-express the remaining integral; if we then further define $u = v^2$ and expand the binomial in the integrand, we get:

$$\int_0^{\pi/2} (\sin\beta)^{2m+1}(\cos\beta)^{2m-\frac{1}{2}} d\beta = 2 \sum_{l=0}^{m} \binom{m}{l} \frac{(-1)^l}{4(l+m)+1}. \tag{B.9}$$

It is perhaps not obvious that this sum can be analytically performed, but in fact it is tractable to computer algebra systems (we used `SymPy` [23]) to get:

$$\int_0^{\pi/2} (\sin\beta)^{2m+1}(\cos\beta)^{2m-\frac{1}{2}} d\beta = \frac{2\Gamma(m+1)\Gamma(m+\frac{5}{4})}{(4m+1)\Gamma(2m+\frac{5}{4})}. \tag{B.10}$$

If we insert this into our expression Eq. (B.8) for $I(x)$, we have now only a single summation and no remaining integrals:

$$I(x) = 8\pi^2 \sum_{m=0}^{\infty} \frac{(\frac{5}{4})_{2m}}{(2m)!} \binom{2m}{m} \frac{\Gamma(m+1)\Gamma(m+\frac{5}{4})}{(4m+1)\Gamma(2m+\frac{5}{4})} x^{2m}. \tag{B.11}$$

Finally, we use the identities $n! = \Gamma(n+1)$, $\Gamma(z+1) = z\Gamma(z)$, and the expression for the binomial coefficient to rewrite this as:

$$I(x) = 8\pi^2 \sum_{m=0}^{\infty} \frac{(\frac{1}{4})_m}{(m)!} (x^2)^m \tag{B.12}$$

which is recognized as the Taylor expansion of $8\pi^2 (1-x^2)^{-1/4}$, as claimed. Given the remarkable simplicity of this result, it would be interesting to find a simpler derivation, but so far we have not.