

Best network chirplet chain: Near-optimal coherent detection of unmodeled gravitational wave chirps with a network of detectors

Archana Pai*

*Max-Planck Institut für Gravitationsphysik, Am Mühlenberg 1, 14476 Potsdam, Germany*Éric Chassande-Mottin[†]*CNRS, AstroParticule et Cosmologie, 10, rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France and Observatoire de la Côte d'Azur, Bd de l'Observatoire, BP 4229, 06304 Nice, France*Olivier Rabaste[‡]*CNRS, AstroParticule et Cosmologie, 10, rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France*

(Received 29 August 2007; published 18 March 2008)

The searches of impulsive gravitational waves (GW) in the data of the ground-based interferometers focus essentially on two types of waveforms: short unmodeled bursts from supernova core collapses and frequency modulated signals (or chirps) from inspiralling compact binaries. There is room for other types of searches based on different models. Our objective is to fill this gap. More specifically, we are interested in GW chirps “in general,” i.e., with an arbitrary phase/frequency vs time evolution. These unmodeled GW chirps may be considered as the generic signature of orbiting or spinning sources. We expect the quasiperiodic nature of the waveform to be preserved independently of the physics which governs the source motion. Several methods have been introduced to address the detection of unmodeled chirps using the data of a single detector. Those include the best chirplet chain (BCC) algorithm introduced by the authors. In the next years, several detectors will be in operation. Improvements can be expected from the joint observation of a GW by multiple detectors and the coherent analysis of their data, namely, a larger sight horizon and the more accurate estimation of the source location and the wave polarization angles. Here, we present an extension of the BCC search to the multiple detector case. This work is based on the coherent analysis scheme proposed in the detection of inspiralling binary chirps. We revisit the derivation of the optimal statistic with a new formalism which allows the adaptation to the detection of unmodeled chirps. The method amounts to searching for salient paths in the combined time-frequency representation of two *synthetic streams*. The latter are time series which combine the data from each detector linearly in such a way that all the GW signatures received are added constructively. We give a proof of principle for the full-sky blind search in a simplified situation which shows that the joint estimation of the source sky location and chirp frequency is possible.

DOI: [10.1103/PhysRevD.77.062005](https://doi.org/10.1103/PhysRevD.77.062005)

PACS numbers: 04.80.Nn, 07.05.Kf, 95.55.Ym

I. SUMMARY

A large effort is underway to analyze the scientific data acquired jointly by the long-baseline interferometric gravitational wave (GW) detectors GEO 600, LIGO, TAMA, and Virgo [1]. In this paper, we contribute to the methodologies employed for this analysis, and, in particular, for the detection of impulsive GW signals.

The current GW data analysis effort is targeted on two types of impulsive GWs. A first target is poorly known short bursts of GWs with a duration in the hundredth of a millisecond range. The astrophysically known sources of such GW bursts are supernovae core collapses (or other similar cataclysmic phenomenon). The second target is frequency modulated signals or *chirps* radiated by inspiralling binaries of compact objects (either neutron stars (NS) or black holes (BH)). These chirp waveforms are

well modeled and expected to last for a few seconds to a few minutes in the detector bandwidth. Our objective is to enlarge the signal range of impulsive GWs under consideration and to “fill the gap” between these two types. More specifically, we are interested in the detection of unmodeled GW chirps which last from a few tens of milliseconds to a few seconds in the detector bandwidth. We shall detail in the next section the astrophysical motivation for considering this kind of GWs.

Joint analysis of the data observed by different GW detectors has obvious benefits. First and foremost, a GW detection can get confirmed or vetoed out with such a joint observation. Further, the detector response depends on the position and orientation of the source and polarization of the wave. For this reason, the joint observation by multiple detectors gives access to physical parameters such as source location and polarization. The use of multiple detectors also allows to enlarge the observational horizon and sky coverage.

Built on the top of pioneering works [2–5], several methods have been proposed and implemented to detect

*Archana.Pai@aei.mpg.de

†Eric.Chassande-Mottin@apc.univ-paris7.fr

‡Olivier.Rabaste@apc.univ-paris7.fr

unmodeled GW chirps which include the signal track search (STS) [6], the chirplet track search [7], and the best chirplet chain (BCC) search proposed by the authors [8]. However, none of the above addresses the multiple detector case. This requires the designing of specific algorithms which are able to combine the information received by the different detectors.

In practice, there are two approaches adopted to carry out network analysis from many detectors: the *coincidence* and *coherent* approaches. In the coincidence approach, the data from each detector is processed independently and only coincident trigger events (in the arrival time and the parameter values) are retained. On the other hand, in the coherent approach, the network as a whole is treated as a single “sensor”: the data from various detectors is analyzed jointly and combined into a single *network statistic* which is tested for detection. In the literature, it has been shown that the coherent approach performs better than the coincidence approach for GW short bursts [9] as well as GW chirps from coalescing binaries [10]. Indeed, the signal phase information is preserved with the coherent approach, whereas it is not with the coincidence approach.

Another reason for this choice is that the coincidence method is not adequate for unmodeled chirps. A large number of parameters (of the same order as the number of signal samples) is needed to characterize their frequency evolution. A coincident detection occurs when the parameter estimates obtained from the analysis of the individual detector data match. Because of the noise perturbation, the occurrence of such a coincidence is very unlikely when the number of parameters is large, unless the incoming GW has very large amplitude. In this article, we adopt the coherent method and propose the *coherent extension of the BCC algorithm*.

Coherent schemes have been already developed for the detection of inspiralling binary chirps [11,12]. Here, we revisit the work presented in [11] with a new formalism. Comments in footnotes link the results presented here with the ones of [11]. We show that the new formalism presented here helps to understand the geometry of the problem and it is simple to establish connections with earlier works.

The outline of the paper goes as follows. In Sec. II, we present and motivate our model of an arbitrary GW chirp. In Sec. III, we describe the response of the detector network to an incoming GW chirp. Further, we show that the linear component of the signal model (parameters acting as scaling factors and phase shifts, so-called extrinsic parameters) can be factorized. This factorization evidences that the signal space can be represented as the direct product of two 2-dimensional spaces i.e., the GW polarization plane and the chirp plane. This representation forms the backbone of the coherent detection scheme that follows in the subsequent section.

In Sec. IV, based on a geometrical argument, we show that the above signal representation manifests the possible

degeneracy of the response. This degeneracy has been already noticed and studied at length in the context of burst detection [13–15]. We investigate this question in the specific context of chirps and obtain similar results as were presented earlier in the literature.

In Sec. V, we obtain the expression of the network statistic. Following the principles of the generalized likelihood ratio test (GLRT), the statistic is obtained by maximizing the network likelihood ratio over the set of unknown parameters. We perform this maximization in two steps. We first treat the linear part of the parametrization and show that such a maximization is nothing but a *least-square* problem over the extrinsic parameters. The solution is obtained by projecting the data onto the signal space. We further study the effect of the response degeneracy on the resulting parameter estimates.

The projection onto the signal space is a combined projection onto the GW polarization and chirp planes. The projection onto the first plane generates two *synthetic streams* which can be viewed as the output of “virtual detectors.” The network statistic maximized over the extrinsic parameters can be conveniently expressed in terms of the processing of those streams. In practice, the synthetic streams linearly combine the data from each detector in such a way that the GW signature received by each detector is added constructively. With this rephrasing, the source location angles can be searched over efficiently.

Along with the projection onto the GW polarization plane, we also examine the projection onto its complement which generates *null streams*. While synthetic streams concentrate the GW contents, the so-called null streams produced this way combine the data such that the GW signal is canceled out. The null streams are useful to veto false triggers due to instrumental artifacts (which do not obey this cancellation property). The null streams we obtained here are identical to the ones presented earlier in GW burst literature [16–18].

In Sec. VI, we perform the second and final step of the maximization of the network statistic over the chirp phase function. This step is the difficult part of the problem. For the one-detector case, we have proposed an efficient method, the BCC algorithm which addresses this question. We show that this scheme can be adapted to the multiple detector case in a straightforward manner, hence we refer to this as *best network CC* (BNCC).

Finally, Sec. VII presents a proof of principle of the proposed method with a full-sky blind search in a simplified situation.

II. GENERIC GW CHIRPS

A. Motivation

Known observable GW sources e.g., stellar binary systems, accreting stellar systems or rotating stars, commonly involve either orbiting or spinning objects. It is not unrea-

sonable to assume that the similar holds true even for the unknown sources.

The GW emission is essentially powered by the source dynamics which thus determines the shape of the emitted waveform. Under linearized gravity and slow motion (i.e., the characteristic velocity is smaller than the speed of light) approximation, the quadrupole formula [19] predicts that the amplitude of the emitted GW is proportional to the second derivative of the quadrupole moment of the physical system. When the dominant part of the bulk motion follows an orbital/rotational motion, the quadrupole moment varies quasiperiodically, and so is the GW.

The more information we have about the GW signal, the better the detection of its signature in the observations. Ideally, this requires precise knowledge of the waveform, and consequently requires precise knowledge of the dynamics. This is not always possible. In general, predicting the dynamics of GW sources in the nearly relativistic regime requires a large amount of effort. This task may get further complicated if mechanisms such as magnetic couplings, mass accretion, density-pressure-entropy gradients, or anisotropic angular momentum distribution are involved.

Here, we are interested in GW sources where the motion is orbital/rotational but the astrophysical dynamics is (totally or partially) unknown. While our primary target is the unforeseen sources (this is why we remain intentionally vague on the exact nature of the sources), several identified candidates enter this category because their dynamics is still not fully characterized. These include (see [8] for more details and references) binary mergers, quasinormal modes from young hot rotating NS, spinning BH accreting from an orbiting disk. As motivated before, following the argument of quadrupole approximation, the GW signature for such sources is not completely undetermined: it is expected to be a quasiperiodic, possibly frequency modulated GW; in brief, it is a *GW chirp*. This is the basic motivation for introducing a generic GW chirp model, as described in the next section.

B. Generic GW chirp model

In this section, we describe the salient features of the generic GW chirp model used in this paper. We motivate the nature of GW polarization, the regularity of its phase, and frequency evolution.

1. Relation between the polarizations

The GW tensor (in the transverse traceless (TT) gauge), associated to the GW emitted from slow-motion, weak gravity sources are mostly due to variations of the mass moments (in contrast to current moments) and can be expanded in terms of mass multipole moments as [20]

$$h^{TT}(t) \propto \sum_{l=2}^{\infty} \sum_{m=-l}^l (\nabla\nabla Y^{lm})^{STF} \frac{d^l}{dt^l} I^{lm}(t - r/c). \quad (1)$$

Here, STF means ‘‘symmetric transverse-tracefree,’’ Y^{lm} are the spherical harmonics, and I^{lm} are the mass multipole moments.

We consider here isolated astrophysical systems with anisotropic mass distributions (e.g., binaries, accreting systems, bar/fragmentation instabilities) orbiting/rotating about a well-defined axis.¹ It can be shown that these systems emit GW predominantly in the $l = |m| = 2$ mode. Contributions from any other mass moments are negligibly small.² The pure-spin tensor harmonic $(\nabla\nabla Y^{lm})^{STF}$ term provides the GW polarization. For $l = |m| = 2$, we have

$$(\nabla\nabla Y^{22})^{STF} \propto (1 + \cos^2\epsilon)e_+ + 2i \cos\epsilon e_{\times}. \quad (2)$$

The tensors e_+ and e_{\times} form a pair of independent and linear-polarization GW tensors (e_{\times} is rotated by $\pi/4$ with respect to e_+). The orbital inclination angle ϵ is the angle between the line of sight to the source (in Earth’s frame) and the angular momentum vector (or the rotation axis) of the physical system, see Fig. 1(a). This shows that the emitted GW in the considered case carries both GW polarizations.

The GW tensor is fully described as $h^{TT}(t) = h_+(t)e_+ + h_{\times}(t)e_{\times}$. The phase shift between the two polarizations h_+ and h_{\times} arises from the I^{lm} term which is proportional to the moment of inertia tensor for $l = |m| = 2$. The quadratic nature of the moment of the inertia tensor introduces a phase shift of $\pi/2$ between the two polarizations h_+ and h_{\times} . This leads to the chirp model below

$$h_+(t) = A \frac{1 + \cos^2\epsilon}{2} \cos(\varphi(t - t_0) + \phi_0), \quad (3a)$$

$$h_{\times}(t) = A \cos\epsilon \sin(\varphi(t - t_0) + \phi_0), \quad (3b)$$

with $t_0 \leq t < t_0 + T$ and $h_{+,\times}(t) = 0$ outside this interval. The phase ϕ_0 is the signal phase at $t = t_0$.

Here, we assume the GW amplitude A to be constant. This is clearly an oversimplified case since we indeed expect an amplitude modulation for real GW sources. However, we wish here to keep the model simple in order to focus the discussion on the aspects related to the coherent analysis of data from multiple detectors. We postpone the study of amplitude modulated GWs to future work.

The chirp model described in Eq. (3) clearly depends upon several unknown parameters (which need to be esti-

¹This condition can be relaxed to precessing systems provided that the precession is over time scales much longer than the observational time, typically of order of seconds.

²Recently, numerical relativity simulations [21] demonstrated that this is a fairly robust statement in the specific context of inspiralling BH binaries. The simulations show that BH binaries emit GW dominantly with $l = |m| = 2$. However, as the mass-ratio decreases, higher multipoles get excited. A similar claim was also made in the context of quasinormal modes produced in the ring-down after the merger of two BH, on the basis of a theoretical argument, see [22], page 4538.

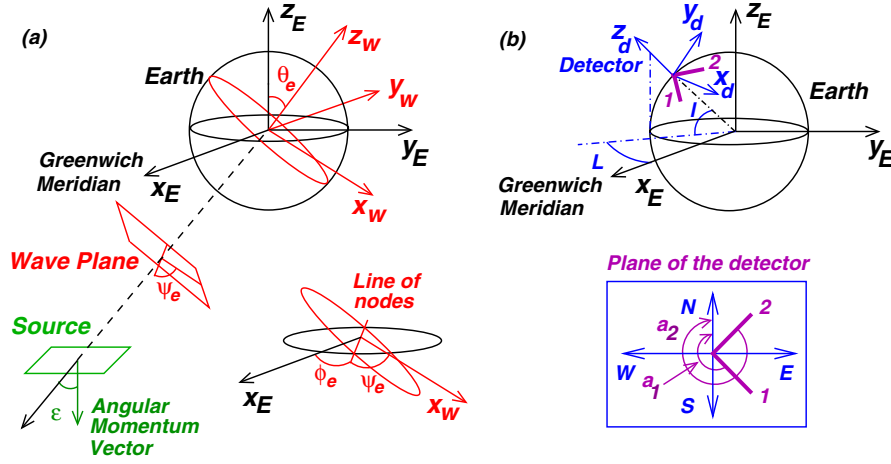


FIG. 1 (color online). *Coordinate transformations.*—(a) $\mathcal{O}(\phi_e, \theta_e, \psi_e)$: Earth frame $\mathbf{x}_E: (x_E, y_E, z_E) \rightarrow$ wave frame $\mathbf{x}_w: (x_w, y_w, z_w)$, (b) $\mathcal{O}(\alpha, \beta, \gamma)$: Earth frame $\mathbf{x}_E: (x_E, y_E, z_E) \rightarrow$ detector frame $\mathbf{x}_d: (x_d, y_d, z_d)$. The latitude l and longitude L of the detector are related to the Euler angles by Eqs. (8) and (9).

mated from the data) which include the amplitude A , the initial phase ϕ_0 , the arrival time t_0 of the chirp, and the inclination angle ϵ . As no precise assumption on the exact nature and dynamics of the GW source is made, we consider the phase evolution function $\varphi(\cdot)$ to be an unknown parameter of the model (3) as well. Clearly, it is a more complicated parameter than the others which are simply scalars. Just like any scalar parameter can be constrained to a range of values (e.g., $A > 0$), the phase function $\varphi(\cdot)$ has to satisfy conditions to be physically realistic which we describe in the next section.

2. Smoothness of the phase evolution

As explained above, the chirp phase is directly related to the orbital phase of the source. The regularity of the orbital phase can be constrained by the physical arguments: the orbital phase and its derivatives are continuous. The same applies to the chirp phase and derivatives.

The detectors operate in a frequency window limited in the range from a few tenths of Hz to a kHz and they are essentially blind outside. This restricts our interest to sources emitting in this frequency range, which results in lower and upper limits on the chirp frequency $\nu(t) \equiv (2\pi)^{-1} d\varphi/dt$ and thus on the variations of the phase.

In addition, the variation of the frequency (the chirping rate) can be connected to the rate at which the source loses its energy. For isolated systems, this rate is clearly bounded. This argument motivates the following bounds on the higher-order derivatives of the phase:

$$\left| \frac{d\nu}{dt} \right| \leq F', \quad \left| \frac{d^2\nu}{dt^2} \right| \leq F''. \quad (4)$$

Equation (4) determines and strengthens the smoothness of the phase/frequency evolution. This is the reason why we coined the term ‘‘smooth GW chirp’’ in [8]. The choice

of the allowed upper bounds F' and F'' may be based on general considerations about the GW source of interest.

We give an example of how those bounds can be set in [8]. We fix F' and F'' according to the variation rate of the frequency of a typical inspiralling binary chirp at the last stable orbit. The resulting chirping rates (estimates obtained from the Newtonian model) can be viewed as a maximum for this kind of system and delimit a sensible range of values.

III. RESPONSE OF A NETWORK OF DETECTORS TO AN INCOMING GW

In this section we derive the response of a network of interferometric ground-based detectors with arbitrary locations and orientations to an incoming GW chirp. The first step is to identify the coordinate frames.

A. Coordinate frames

We follow the conventions of [11] and introduce three coordinate frames, namely, the wave frame, the Earth frame, and the detector frame as given below, see Fig. 1.

- (i) the *wave* frame $\mathbf{x}_w \equiv (x_w, y_w, z_w)$ is the frame associated to the incoming GW with positive z_w -direction along the incoming direction and the $x_w - y_w$ plane corresponds to the plane of the polarization of the wave.
- (ii) the *Earth* frame $\mathbf{x}_E \equiv (x_E, y_E, z_E)$ is the frame attached to the center of the Earth. The x_E axis is radially pointing outwards from the Earth’s center and the equatorial point that lies on the meridian passing through Greenwich, England. The z_E axis points radially outwards from the center of Earth to the North Pole. The y_E axis is chosen to form a right-handed coordinate system with the x_E and z_E axes.

- (iii) the *detector* frame $\mathbf{x}_d \equiv (x_d, y_d, z_d)$ is the frame attached to the individual detector. The $(x_d - y_d)$ plane contains the detector arms and is assumed to be tangent to the surface of the Earth. The x_d axis bisects the angle between the detector's arms. The z_d axis points towards the local zenith. The direction of the y_d axis is chosen so that we get a right-handed coordinate system.

A rotation transformation between the coordinate systems about the origin is specified by the rotation operator \mathcal{O} which is characterized by three Euler angles. We define these angles using the “ x -convention” (also known as $z - x - z$ convention) [23].

Let $(\phi_e, \theta_e, \psi_e)$ and (α, β, γ) be Euler angles of the rotation operator relating pairs of the above coordinate systems as follows:

$$\mathbf{x}_w = \mathcal{O}(\phi_e, \theta_e, \psi_e)\mathbf{x}_E, \quad (5)$$

$$\mathbf{x}_d = \mathcal{O}(\alpha, \beta, \gamma)\mathbf{x}_E. \quad (6)$$

All the angles in Eqs. (5) and (6) are related to physical/geometrical quantities described in Fig. 1. More specifically, we have

$$\phi_e = \phi - \pi/2, \quad \theta_e = \pi - \theta, \quad \psi_e = \psi, \quad (7)$$

where ϕ and θ are the spherical polar coordinates of the source in the Earth's frame and the angle ψ is the so-called polarization-ellipse angle which gives the orientation of the source plane. Throughout the paper, we shall use θ and ϕ to indicate the source location.

The detector Euler angles (α, β, γ) are directly related to the location and orientation of the detector as follows:

$$\alpha = L + \pi/2, \quad (8)$$

$$\beta = \pi/2 - l, \quad (9)$$

$$\gamma = \frac{a_1 + a_2}{2} + \frac{3\pi}{2} \quad \text{if } |a_1 - a_2| > \pi, \quad (10)$$

$$= \frac{a_1 + a_2}{2} + \frac{\pi}{2} \quad \text{if } |a_1 - a_2| \leq \pi, \quad (11)$$

where l and L are the latitude and longitude of the corner station. The angles a_1 and a_2 describe the orientation of the first and second arm, respectively. It is the angle through which one must rotate the arm clockwise (while viewing from top) to point the local north. In Table I, we tabulate the currently running interferometric detectors along with their Euler angles.

Combining Eqs. (5) and (6), we obtain the coordinate transformation from the wave frame to the detector frame as follows:

$$\mathbf{x}_s = \mathcal{O}(\phi'_e, \theta'_e, \psi'_e)\mathbf{x}_d, \quad (12)$$

where $\mathcal{O}(\phi'_e, \theta'_e, \psi'_e) \equiv \mathcal{O}(\phi_e, \theta_e, \psi_e)\mathcal{O}^{-1}(\alpha, \beta, \gamma)$.

B. Network response

The detector response to an incident GW is obtained by contracting the GW tensor with the detector tensor [see Appendix B], which can be reexpressed as a linear combination of the two polarizations h_+ and h_\times i.e.

$$s = f_+ h_+ + f_\times h_\times \equiv \Re[f^* h]. \quad (13)$$

The linear coefficients $f_+(\phi'_e, \theta'_e, \psi'_e)$ and $f_\times(\phi'_e, \theta'_e, \psi'_e)$, commonly termed as the detector antenna pattern functions, represent the detector's directional response to the $+$ and \times polarizations, respectively. For the compact expression provided by Eq. (13), we have defined the complex GW signal to be $h = h_+ + ih_\times$ and the complex antenna pattern function to be $f = f_+ + if_\times$.

The detector response s and the incident GW signal h are both times series. In a network where the various detectors are located at different locations on the Earth, for example, the LIGO-Virgo network, the GW arrives at the detector sites at different time instances. However, all the measurements at the various detectors need to be carried out with a reference time. Here, for our convenience, the time measured by an observer attached to the Earth's center as a reference is treated as a reference. Any other reference would be equally acceptable. We have

TABLE I. *Location and orientation of the GW detectors.*—We identify the detectors with two character label. Concerning the LIGO detectors, H1 and L1 refer to the Hanford (WA) and Livingston (LA) detectors, respectively. The location of the corner station (vertex) of each detector is given in terms of the latitude and longitude. The longitudes with an East (respectively West) suffix receive a “+” sign (respectively “−” sign) when converted into radians in Eq. (8). The orientation of the arm is given by the angle through which one must rotate it clockwise (while viewing from top) to point the local north. The corresponding detector Euler angles (α, β, γ) are listed.

Detector name & label	Vertex position		Arm orientation		Euler angles		
	latitude (N)	longitude	a_1	a_2	α	β	γ
TAMA (T1)	35° 40' 35.60''	139° 32' 09.80'' E	90° 00' 00.00''	179° 59' 60.00''	229° 32' 09.80''	54° 19' 24.40''	225° 00' 00.00''
GEO (G1)	52° 14' 42.53''	9° 48' 25.89'' E	291° 36' 42.12''	25° 56' 35.16''	99° 48' 25.89''	37° 45' 17.47''	68° 46' 38.64''
Virgo (V1)	43° 37' 53.09''	10° 30' 16.19'' E	340° 34' 02.03''	70° 34' 02.03''	100° 30' 16.19''	46° 22' 06.91''	115° 34' 02.03''
LIGO (H1)	46° 27' 18.53''	119° 24' 27.57'' W	35° 59' 57.84''	125° 59' 57.84''	−29° 24' 27.57''	43° 32' 41.47''	170° 59' 57.84''
LIGO (L1)	30° 33' 46.42''	90° 46' 27.27'' W	107° 42' 59.40''	197° 42' 59.40''	−0° 46' 27.27''	59° 26' 13.58''	242° 42' 59.40''

$$s(t) = \Re[f^* h(t - \tau(\phi, \theta))], \quad (14)$$

where $\tau(\phi, \theta) = (\mathbf{r}_d - \mathbf{r}_E) \cdot \mathbf{w}(\phi, \theta)/c$ denotes the difference in the arrival times of the GW (propagating with the unit wave vector \mathbf{w}) at the detector and at the center of the Earth located at \mathbf{r}_d and \mathbf{r}_E , respectively. Note that this value can be positive or negative depending on the source location.

C. Vector formalism

In the following, we distinguish scalars by using roman letters, vectors are denoted by small bold letters, and matrices by bold capitals. We denote the k -th element of vector \mathbf{a} by $a[k]$ and correspondingly, the element of matrix \mathbf{A} at row k and column l by $A[k, l]$. The matrices \mathbf{A}^T and $\mathbf{A}^H \equiv (\mathbf{A}^T)^*$ designate the real and Hermitian transposes of \mathbf{A} , respectively.

We consider now a GW detector network with d interferometers. Each detector and its associated quantities are labeled with an index $j = 1, \dots, d$ which we also use as a subscript if required. We assume that the output response of each detector is sampled at the Nyquist rate $\nu_s \equiv 1/t_s$ where t_s is the sampling interval. We then divide the data in blocks of N consecutive samples. In this setup, the detector as well as the network response is then defined by forming vectors with these blocks of data.

Let us consider a given GW chirp source at sky location (ϕ, θ) . Let the response of the j -th detector be \mathbf{s}_j with entry $s_j[k] = s_j(t_k + \tau_j(\phi, \theta))$, where $t_k = t_0 + kt_s$, $k = 0, \dots, N-1$, and t_0 is the reference time i.e. the time of arrival of GW at the center of the Earth. Note that, with the above definition, we compensate for the time delay $\tau_j(\phi, \theta)$ between the detector j and the Earth's center. Thus, in this setup, the GW signal starts and ends in the same rows in the data vectors \mathbf{s}_j of all the detectors.

For compactness, we stack the data from all the detectors in the network into a single vector \mathbf{s} of size $Nd \times 1$, such that $\mathbf{s}^T = [\mathbf{s}_1^T \mathbf{s}_2^T \dots \mathbf{s}_d^T]$ forms the *network response*. In this convention, the network response can be expressed compactly as the *Kronecker product* (see Appendix A for the definition) of the network complex beam pattern vector $\mathbf{f} = \{f_j, j = 1 \dots d\} \in \mathbb{C}^{d \times 1}$ and the complex GW vector $\mathbf{h} = \{h(t_k), t_k = t_0 + t_s k \text{ with } k = 0 \dots N-1\} \in \mathbb{C}^{N \times 1}$ viz.,

$$\mathbf{s} = \Re[\mathbf{f}^* \otimes \mathbf{h}]. \quad (15)$$

The above expression is general enough to hold true for any type of incoming GW signal. The Kronecker product in this expression is the direct manifestation of the fact that the detector response is nothing but the tensor product between the detector and the wave tensors.

D. GW chirp as a linear model of the extrinsic parameters

In the previous section, we have obtained the network response to any type of incoming GW with two polar-

izations. In what follows, we wish to investigate how this response manifests in the case of a specific type of GW, namely, GW chirp described in Eq. (3). We also want to understand how various parameters explicitly appear in the network response.

It is insightful to distinguish the signal parameters based on their effect on the signal model. The parameters are separated into two distinct types traditionally referred to as the “*intrinsic*” and “*extrinsic*” parameters. The extrinsic parameters are those that introduce scaling factors or phase shifts but do not affect the shape of the signal model. Instead, intrinsic parameters significantly alter the shape of the signal and hence the underlying geometry.

The network response \mathbf{s} mingles these two types of parameters. Our work is considerably simplified if we can “factorize” the extrinsic parameters from the rest. For the chirp model described in Eq. (3), we count four extrinsic parameters, namely $\{A, \phi_0, \epsilon, \psi_e\}$ and perform this factorization in two steps.

1. Extended antenna pattern includes the inclination angle

We absorb the inclination angle ϵ into the antenna pattern functions and rewrite the network signal as

$$\mathbf{s} = \Re[\tilde{\mathbf{f}}^* \otimes \tilde{\mathbf{h}}], \quad (16)$$

where $\tilde{\mathbf{h}} \equiv a\mathbf{e}$ is the GW vector. It only depends on the complex amplitude $a = A \exp(i\phi_0)$ and on the phase vector $\mathbf{e} = \{\exp(i\varphi[k]), \text{ with } \varphi[k] \equiv \varphi(kt_s), k = 0 \dots N-1\}$.

The *extended* antenna pattern $\tilde{\mathbf{f}}$ incorporates the inclination angle ϵ as follows³

$$\tilde{\mathbf{f}} = \frac{1 + \cos^2 \epsilon}{2} \mathbf{f}_+ + i \cos \epsilon \mathbf{f}_\times. \quad (17)$$

2. Gel'fand functions factorize the polarization angles from source location angles

The second step is to separate the dependency of $\tilde{\mathbf{f}}$ on the polarization angles $\{\psi, \epsilon\}$ from the source location angle and the detector orientation angles. The earlier work [11] shows that the Gel'fand functions (which are a representation of the rotation group $SO(3)$) provide an efficient tool to do the same. For the sake of completeness, Appendix B reproduces some of the calculations of [11]. The final result (see also Eqs. (3.14–3.16) of [11]) yields the following decomposition:

$$\tilde{\mathbf{f}} = t_+ \mathbf{d} + t_- \mathbf{d}^*, \quad (18)$$

where the vector $\mathbf{d} \in \mathbb{C}^{d \times 1}$ carries the information of the source location angles (ϕ, θ) via (ϕ_e, θ_e) and the detector Euler angles $\{\alpha_j, \beta_j, \gamma_j\}$. Its components are expressed as

³We remind the reader that a similar quantity was previously introduced in Eq. (3.19) of [11].

$$d[j] = - \sum_{n=-2}^{n=2} iT_{2n}(\phi_e, \theta_e, 0) \times [T_{2n}(\alpha_j, \beta_j, \gamma_j) - T_{-2n}(\alpha_j, \beta_j, \gamma_j)]^*. \quad (19)$$

The tensor T_{mn} designates rank-2 Gel'fand functions. The coefficients t_+ and t_- depend only on the polarization angles $\{\psi, \epsilon\}$, viz.

$$t_{\pm} = T_{2\pm 2}(\psi, \epsilon, 0) = \frac{(1 \pm \cos \epsilon)^2}{4} \exp(\mp 2i\psi). \quad (20)$$

Finally, we combine Eqs. (16) and (18) and obtain an expression of the network response where the extrinsic parameters are ‘‘factorized’’ as follows:

$$\mathbf{s} = \left(\underbrace{[\mathbf{d} \ \mathbf{d}^*]}_{\mathbf{D}} \otimes \frac{1}{2} \underbrace{[\mathbf{e} \ \mathbf{e}^*]}_{\mathbf{E}} \right) \underbrace{\begin{pmatrix} at_-^* \\ a^* t_+ \\ at_+^* \\ a^* t_- \end{pmatrix}}_{\mathbf{p}} \equiv \mathbf{\Pi} \mathbf{p}. \quad (21)$$

Equation (21) evidences the underlying linearity of the GW model with respect to the extrinsic parameters. The 4-dimensional complex vector \mathbf{p} defines a one-to-one (non-linear) mapping between its components and the four physical extrinsic parameters $\{A, \phi_0, \epsilon, \psi\}$ (we will detail this point later in Sec. VA 3). Note that the first and fourth components as well as the second and third components of \mathbf{p} are complex conjugates. This symmetry comes from the fact that the data is real.

The *signal space* as defined by the network response is the range of $\mathbf{\Pi}$ and results from the Kronecker product of two linear spaces: the plane of \mathbb{C}^d generated by the columns of \mathbf{D} which we shall refer to as *GW polarization plane*⁴ and the plane of \mathbb{C}^N generated by the columns of \mathbf{E} which we shall refer to as *chirp plane*. These two spaces embody two fundamental characteristics of the signal: the former characterizes gravitational waves while the latter characterizes chirping signals. The Kronecker product in the expression of $\mathbf{\Pi}$ shows explicitly that the network response is the result of the *projection* of incoming GW onto the detector network.

The norm of the network signal gives the ‘‘signal’’ (and not physical) energy delivered to the network, which is

$$\|\mathbf{s}\|^2 = \frac{NA^2}{2} \|\tilde{\mathbf{f}}\|^2. \quad (22)$$

Clearly, the dependence on the number of samples N implies that the longer the signal duration, the larger the signal energy and is proportional to the length of the signal duration. The factor $\|\tilde{\mathbf{f}}\|$ is the modulus of the extended antenna pattern vector. It can be interpreted as the gain or

⁴In [11], this plane was referred to as ‘‘helicity plane’’ because it is formed by the network beam patterns for all possible polarizations.

attenuation depending on the direction of the source and on the polarization of the wave.

IV. INTERPRETATION OF THE NETWORK RESPONSE

In this section, we focus on understanding the underlying geometry of the signal model described in Eq. (21). A useful tool to do so is the singular value decomposition (SVD) [24]. It provides an insight on the geometry by identifying the principal directions of linear transforms.

A. Principal directions of the signal space: Singular value decomposition

The SVD is a generalization of the eigen-decomposition for nonsquare matrices. The SVD factorizes a matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ into a product $\mathbf{A} = \mathbf{U}_A \mathbf{\Sigma}_A \mathbf{V}_A^H$ of three matrices $\mathbf{U}_A \in \mathbb{C}^{m \times r}$, $\mathbf{\Sigma}_A \in \mathbb{R}^{r \times r}$, and $\mathbf{V}_A \in \mathbb{C}^{n \times r}$ where $r \leq m, n$ is the rank of \mathbf{A} . The columns of \mathbf{U}_A and \mathbf{V}_A are orthonormal i.e., $\mathbf{U}_A^H \mathbf{U}_A = \mathbf{V}_A^H \mathbf{V}_A = \mathbf{I}_r$. The diagonal of $\mathbf{\Sigma}_A$ are the singular values (SV) of \mathbf{A} . We use here the so-called ‘‘compact’’ SVD (we retain the nonzero SV only in the decomposition), such that the matrix $\mathbf{\Sigma}_A$ is a positive definite diagonal matrix.

The SVD is compatible with the Kronecker product [25]: the SVD of a Kronecker product is the Kronecker product of the SVDs. Applying this property to $\mathbf{\Pi}$, we get

$$\mathbf{\Pi} = (\mathbf{U}_D \otimes \mathbf{U}_E) (\mathbf{\Sigma}_D \otimes \mathbf{\Sigma}_E) (\mathbf{V}_D \otimes \mathbf{V}_E)^H. \quad (23)$$

Therefore, the SVD of $\mathbf{\Pi}$ can be easily deduced from the one of \mathbf{D} and \mathbf{E} . We note that \mathbf{D} and \mathbf{E} have similar structure (two complex conjugated columns), see Eq. (21). In Appendix C, we analytically obtain the SVD of a matrix with such a structure. Thus, applying this result, we can straightaway write down the SVDs for \mathbf{D} and \mathbf{E} as shown in the following sections.

1. GW polarization plane: SVD of \mathbf{D}

Let us first introduce some variables

$$\mathcal{D} \equiv \mathbf{d}^H \mathbf{d} = \sum_{j=1}^d |d[j]|^2, \quad (24)$$

$$\Delta \equiv \mathbf{d}^T \mathbf{d} = \sum_{j=1}^d d[j]^2, \quad (25)$$

$$\delta \equiv \arg \Delta. \quad (26)$$

In the nominal case, the matrix \mathbf{D} has rank 2, viz.

$$\mathbf{\Sigma}_D = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}, \quad (27)$$

with two nonzero SV $\sigma_1 = \sqrt{\mathcal{D} + |\Delta|}$ and $\sigma_2 = \sqrt{\mathcal{D} - |\Delta|}$ ($\sigma_1 \geq \sigma_2$) associated to a pair of left-singular

vectors $\mathbf{V}_D = [\mathbf{v}_1, \mathbf{v}_2]$ with

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} \exp(-i\delta) \\ 1 \end{bmatrix}, \quad \mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} \exp(-i\delta) \\ -1 \end{bmatrix}, \quad (28)$$

and of right-singular vectors $\mathbf{U}_D = [\mathbf{u}_1, \mathbf{u}_2]$ with

$$\mathbf{u}_1 = \frac{\exp(-i\delta)\mathbf{d} + \mathbf{d}^*}{\sqrt{2(\mathcal{D} + |\Delta|)}}, \quad (29)$$

$$\mathbf{u}_2 = \frac{\exp(-i\delta)\mathbf{d} - \mathbf{d}^*}{\sqrt{2(\mathcal{D} - |\Delta|)}}. \quad (30)$$

Note that the vector pair $\{\mathbf{u}_1, \mathbf{u}_2\}$ results from the Gram-Schmidt orthonormalization of $\{\mathbf{d}, \mathbf{d}^*\}$.

Barring the nominal case, for a typical network built with the existing detectors and for certain sky locations of the source, it is, however, possible for the smallest SV σ_2 to vanish. In such a situation, the rank of \mathbf{D} reduces to 1. We then have $\Sigma_D = \sigma_1$, $\mathbf{V}_D = \mathbf{v}_1$ and $\mathbf{U}_D = \mathbf{u}_1$. We give an interpretation of this degeneracy later in Sec. IV B.

2. Chirp plane: SVD of \mathbf{E}

The results of the previous section essentially apply to SVD calculation of \mathbf{E} . However, there is an additional simplification due to the nature of the columns of \mathbf{E} . Indeed, the cross-product

$$\mathbf{e}^T \mathbf{e} = \sum_{k=0}^{N-1} \exp(2i\varphi[k]), \quad (31)$$

is an oscillating sum. This sum can be shown [8] to be of small amplitude under mild conditions compatible with the case of interest. We can thus consider⁵ that $\mathbf{e}^T \mathbf{e} \approx 0$ and $\mathbf{e}^H \mathbf{e} = N$. Therefore, following Appendix C 2, the SVD of \mathbf{E} is given by $\Sigma_E = \sqrt{N}\mathbf{I}_2/2$, $\mathbf{V}_E = \mathbf{I}_2$, and $\mathbf{U}_E = 2\mathbf{E}/\sqrt{N}$.

3. Signal space: SVD of $\mathbf{\Pi}$

We obtain the SVD for $\mathbf{\Pi}$ using the compatibility of the SVD with the Kronecker product stated in Eq. (23). In the nominal case where \mathbf{D} has rank 2, we have

$$\Sigma_{\Pi} = \frac{\sqrt{N}}{2} \begin{bmatrix} \sigma_1 \mathbf{I}_2 & \mathbf{0}_2 \\ \mathbf{0}_2 & \sigma_2 \mathbf{I}_2 \end{bmatrix}, \quad (32)$$

with four left-singular vectors

$$\mathbf{V}_{\Pi} = [\mathbf{v}_1 \otimes \mathbf{I}_2 \quad \mathbf{v}_2 \otimes \mathbf{I}_2], \quad (33)$$

and four right-singular vectors

$$\mathbf{U}_{\Pi} = \frac{2}{\sqrt{N}} [\mathbf{u}_1 \otimes \mathbf{e} \quad \mathbf{u}_1 \otimes \mathbf{e}^* \quad \mathbf{u}_2 \otimes \mathbf{e} \quad \mathbf{u}_2 \otimes \mathbf{e}^*]. \quad (34)$$

B. The signal model can be ill-posed

In the previous section, we obtained the SVD of $\mathbf{\Pi}$ in the nominal case where the matrix \mathbf{D} has 2 nonzero SVs. As we have already mentioned, for a typical detector network, there might exist certain sky locations where the second SV σ_2 of \mathbf{D} vanishes which implies that the rank of \mathbf{D} degenerates to 1. In such cases, this degeneracy propagates to $\mathbf{\Pi}$ and subsequently its rank reduces from 4 to 2.

In order to realize the consequences of this degeneracy, we first consider a network of *ideal* GW detectors (with no instrumental noise). Let a GW chirp pass through such a network from a source in a sky location where $\sigma_2 = 0$. The detector output is *exactly* equal to \mathbf{s} . An estimate of the source parameters would then be obtained from the network data by inverting Eq. (21). However, in this case, this is impossible since it requires the inversion of an under-determined linear system (there are 4 unknowns and only 2 equations).

This problem is identical to the one identified and discussed at length in a series of articles devoted to unmodeled GW bursts [13–15], where this problem is formulated as follows: at those sky locations where \mathbf{D} is degenerated, the GW response is essentially made of only *one* linear combination of the *two* GW polarizations. It is thus impossible to separate the two individual polarizations (unless additional prior information is provided). We want to stress here that this problem is not restricted to unmodeled GW bursts but also affects the case of chirping signals (and extends to the chirps from inspiralling binaries of NS or BH⁶). This is mainly because the degeneracy arises from the geometry of the GW polarization plane which is the same for any type of source.

The degeneracy disappears at locations where $\sigma_2 > 0$ even if it is infinitesimally small. However, when σ_2 is small, the inversion of the linear equations in Eq. (21) is very sensitive to perturbations. With *real-world* GW detectors, instrumental noise affects the detector response i.e., perturbs the left-hand side of Eq. (21).

A useful tool to investigate this is the *condition number* [15]. It is a well-known measure of the sensitivity of linear systems. The condition number of a matrix \mathbf{A} is defined as the ratio of its largest SV to the smallest. For unitary matrices, $\text{cond}(\mathbf{A}) = 1$. On the contrary, if \mathbf{A} is rank deficient, $\text{cond}(\mathbf{A}) \rightarrow \infty$. For the matrix $\mathbf{\Pi}$, we have

⁵This amounts to saying that the two GW polarizations (i.e., the real and imaginary parts of $\exp i\varphi[k]$) are orthogonal and of equal norm. Note that this approximation is not required and can be relaxed. This would lead to use a version of the polarization pair orthonormalized with a Gram-Schmidt procedure.

⁶Contrary to the generic chirp model considered here, the phase and amplitude functions of inspiralling binary chirps follow a prescribed power-law time evolution. These differences affect only the geometry of the “chirp plane,” but not that of the “GW polarization plane,” hence the conclusion on the degeneracy remains the same.

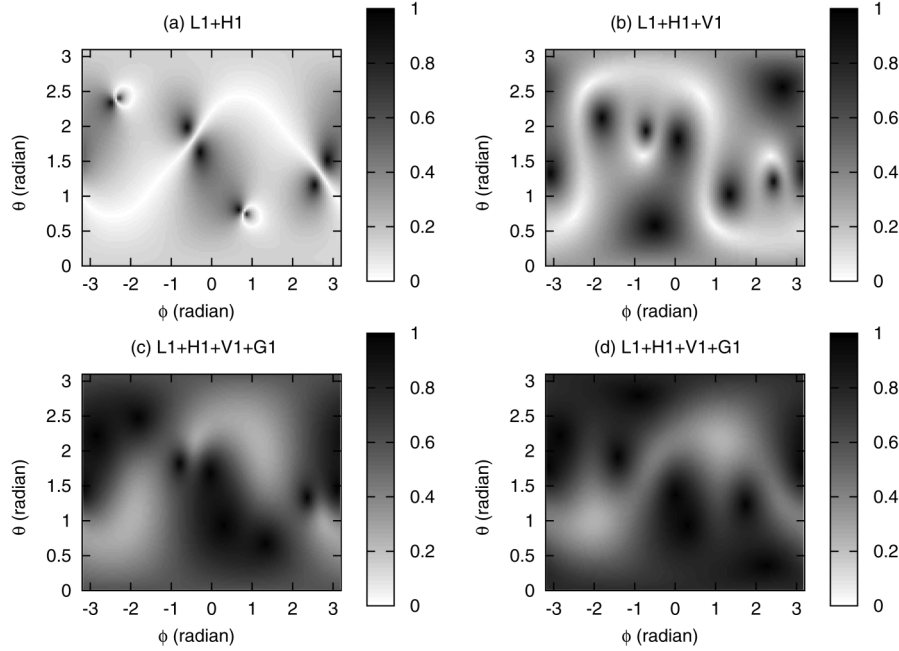


FIG. 2. *Degeneracy of the network response.*—We show here the inverse of $\text{cond}(\mathbf{\Pi})$ for various detector networks (the abbreviated detector names are listed in Table I). The brighter regions of the sky correspond to the large conditioning number $\text{cond}(\mathbf{\Pi})$. The fraction of the sky where $1/\text{cond}(\mathbf{\Pi}) > 0.1$ is (a) 27%, (b) 4.8%, (c) and (d) 0%. Since the LIGO detectors are almost aligned, this pair of detectors has the largest percentage of degeneracy.

$$\text{cond}(\mathbf{\Pi}) = \frac{\sigma_1}{\sigma_2} = \sqrt{\frac{\mathcal{D} + |\Delta|}{\mathcal{D} - |\Delta|}}. \quad (35)$$

In Fig. 2, we show full-sky plots of $1/\text{cond}(\mathbf{\Pi})$ for various configurations (these figures essentially reproduce the ones of [15]). We see that, even for networks of misaligned detectors, there are significantly large patches where $\text{cond}(\mathbf{\Pi})$ takes large values. In those regions, the inversion of Eq. (21) is sensitive to the presence of noise and the estimate of the extrinsic parameters thus have a large variance.

a. Connection to the antenna pattern function—Interestingly, the SVs of \mathbf{D} and $\mathbf{F} \equiv [\mathbf{f}\mathbf{f}^*]$ coincide. This can be seen from the following relationships we directly obtained from the definitions in Eqs. (17), (18), and (20)

$$\begin{bmatrix} \tilde{\mathbf{f}} & \tilde{\mathbf{f}}^* \end{bmatrix} = \mathbf{D} \begin{bmatrix} t_+ & t_-^* \\ t_- & t_+^* \end{bmatrix} = \mathbf{F} \begin{bmatrix} |t_+| & |t_-| \\ |t_-| & |t_+| \end{bmatrix}. \quad (36)$$

The matrix \mathbf{F} can be obtained from \mathbf{D} by a unitary transformation. Both matrices share the same singular spectrum. We can therefore write

$$\sigma_2^2 = \sum_{j=1}^d |f[j]|^2 - \left| \sum_{j=1}^d f[j] \right|^2. \quad (37)$$

When $\sigma_2 \sim 0$, we thus have $|\mathbf{f}_+ \times \mathbf{f}_\times|^2 \sim 0$ where $\mathbf{f}_+ \equiv \Re[\mathbf{f}]$ and $\mathbf{f}_\times \equiv \Im[\mathbf{f}]$ are the network antenna pattern vectors. This means that at such sky locations, the antenna pattern vectors get aligned even if the detectors in the

network are misaligned. In other words, despite that the considered network is composed of misaligned detectors, it acts as a network of aligned detectors at those sky locations. (Of course, for perfectly coaligned detectors, $\mathbf{f}_+ \propto \mathbf{f}_\times$ at all sky locations.) Networks with many detectors having different orientations are less likely to be degenerate. This is confirmed in Fig. 2 where we see that the size of the degenerate sky patches reduces rapidly when the number of detector with varied orientations increases. The network formed by the four detectors L1-H1-V1-G1 (assuming they have the same noise spectrum) does not show any patches with significant degeneracy.

V. NETWORK LIKELIHOOD ANALYSIS: GW POLARIZATION PLANE AND SYNTHETIC STREAMS

Generally speaking, a signal detection problem amounts to testing the null hypothesis (H_0) (absence of signal in the data) vs the alternate hypothesis (H_1) (presence of signal in the data). Because of the presence of noise, two types of errors occur: false dismissals (decide H_0 when H_1 is present) and false alarms (decide H_1 when H_0 is present). There exist several objective criteria to determine the detection procedure (or statistic) which optimizes the occurrence of these errors. We choose the Neyman-Pearson (NP) approach which minimizes the number of false dismissals for a fixed false alarm rate. It is easily shown that for simple problems, the likelihood ratio (LR) is NP optimal. However, when the signal depends upon unknown

parameters, the NP optimal (uniformly overall allowed parameter values) statistic is not easy to obtain. Indeed for most real-world problems, it does not even exist. However, the generalized likelihood ratio test (GLRT) [26] has shown to give sensible results and hence is widely used. In the GLRT approach, the parameters are replaced by their maximum likelihood estimates. In other words, the GLRT approach uses the maximum likelihood ratio as the statistic. Here, we opt for such a solution.

As a first step, we consider the simplified situation where all detectors have independent and identical instrumental noises and this noise is white and Gaussian with unit variance. We will address the colored noise case later in Sec. VD.

In this case, the logarithm of the network likelihood ratio (LLR) is given by

$$\Lambda(\mathbf{x}) = -\|\mathbf{x} - \mathbf{s}\|^2 + \|\mathbf{x}\|^2, \quad (38)$$

where $\|\cdot\|^2$ is the Euclidean norm (here in \mathbb{R}^{Nd}) and we omitted an unimportant factor $1/2$. The network data vector \mathbf{x} is constructed on similar lines as that of the network response \mathbf{s} , i.e. first, it stacks the data from all the detectors into $\mathbf{x}^T = [\mathbf{x}_1^T, \mathbf{x}_2^T, \dots, \mathbf{x}_d^T]$ and then at each detector, the data is time-shifted to account for the delay in the arrival time $\mathbf{x}_j = \{x_j[k] = x_j(t_k + \tau_j), t_k = t_s k \text{ and } k = 0 \dots N - 1\}$.

A. Maximization over extrinsic parameters: scaling factors and phase shifts

Following the GLRT approach, we maximize the network LLR Λ with respect to the parameters of \mathbf{s} . We replace \mathbf{s} by its model as given in Eq. (21) and consider at first the maximization with respect to the extrinsic parameters \mathbf{p} .

1. Least-square fit

The maximization of the network LLR over \mathbf{p} amounts to fitting a linear signal model to the data in *least-square* (LS) sense, viz.

$$\text{minimize } -\Lambda(\mathbf{x}) + \|\mathbf{x}\|^2 = \|\mathbf{x} - \Pi\mathbf{p}\|^2 \text{ over } \mathbf{p}. \quad (39)$$

This LS problem is easily solved using the pseudoinverse $\Pi^\#$ of Π [24]. The estimate of \mathbf{p} is then given by

$$\hat{\mathbf{p}} = \Pi^\# \mathbf{x}. \quad (40)$$

The pseudoinverse can be expressed using the SVD of Π as $\Pi^\# = \mathbf{V}_\Pi \Sigma_\Pi^{-1} \mathbf{U}_\Pi^H$ (note that $\Pi^\#$ is always defined since we use the *compact* SVD restricted to nonzero SVs).

Substituting Eq. (40) in Eq. (39), we get the LS minimum to be

$$-\hat{\Lambda}(\mathbf{x}) + \|\mathbf{x}\|^2 = \|\mathbf{x} - \mathbf{U}_\Pi \mathbf{U}_\Pi^H \mathbf{x}\|^2, \quad (41)$$

where we used $\mathbf{V}_\Pi^H \mathbf{V}_\Pi = \mathbf{I}_r$. Equation (41) can be further simplified into⁷

$$\hat{\Lambda}(\mathbf{x}) = \|\mathbf{U}_\Pi^H \mathbf{x}\|^2. \quad (42)$$

It is interesting to note that the operator $\mathbf{U}_\Pi \mathbf{U}_\Pi^H$ is a (orthogonal) projection operator onto the signal space (over the range of Π) i.e. $\mathbf{U}_\Pi \mathbf{U}_\Pi^H \Pi = (\Pi \Pi^\#) \Pi = \Pi$.

2. Signal-to-noise ratio

The signal-to-noise ratio (SNR) measures the difficulty level for detecting a signal in the noise. In the present case, along with the amplitude and duration of the incoming GW, the network SNR also depends on the relative position, orientation of the source with respect to the network. Therefore, the SNR should incorporate all these aspects. A systematic way to define the SNR is to start from the statistic.

Let the SNR ρ of an injected GW chirp $\mathbf{s}_0 = \Pi \mathbf{p}_0$ be⁸

$$\rho^2 \equiv \hat{\Lambda}(\mathbf{s}_0). \quad (43)$$

Note that in this expression, the matrix Π in the statistic and in \mathbf{s}_0 are the same. Using the SVD of Π and the property of the projection operator $\mathbf{U}_\Pi^H \mathbf{U}_\Pi = \mathbf{I}_r$, we get $\rho^2 = \|\mathbf{s}_0\|^2$. The SNR is equal to the ‘‘signal energy’’ in the network data as defined in Eq. (22). Thus, the SNR ρ scales as \sqrt{N} as expected and it depends on the source direction, polarization, and network configuration through the gain factor $\|\tilde{\mathbf{f}}\|$.⁹ Figure 3 illustrates how this factor varies for the network formed by the two LIGO detectors and Virgo. Figure 3 displays the ratio ρ/ρ_{best} between the global SNR (obtained with a coherent analysis) and the largest individual SNR (obtained with the best detector of the network). The panels (a) and (b) are associated to the ‘‘worst’’ (minimum over all polarization angles ϵ and ψ) and ‘‘best’’ (maximum) cases, respectively. Ideally, when the detectors are aligned, the enhancement factor is expected to be \sqrt{d} (≈ 1.73 in the present case). In the best case, the enhancement is $\gtrsim 1.7$ for more than half of the sky (94% of the sky when $\gtrsim 1.4$). In the worst case, the SNR enhancement is 1.28 at most and 8.5% of the sky gets a value $\gtrsim 1.1$.

3. From geometrical to physical parameter estimates

The components of \mathbf{p} do not have a direct physical interpretation but as mentioned earlier, they are rather functions of the physical parameters. Following the above

⁷For the inspiral case, this expression is equivalent to Eq. (4.8) of [11].

⁸If the noise power is not unity, it would divide the signal energy in this expression. When we have only one detector, the SNR ρ^2 is consistent with the definition usually adopted in this case.

⁹The SNR ρ^2 is similar to b^2 defined in Eq. (3.17) of [11] in the case of inspiralling binary signal and colored noise.

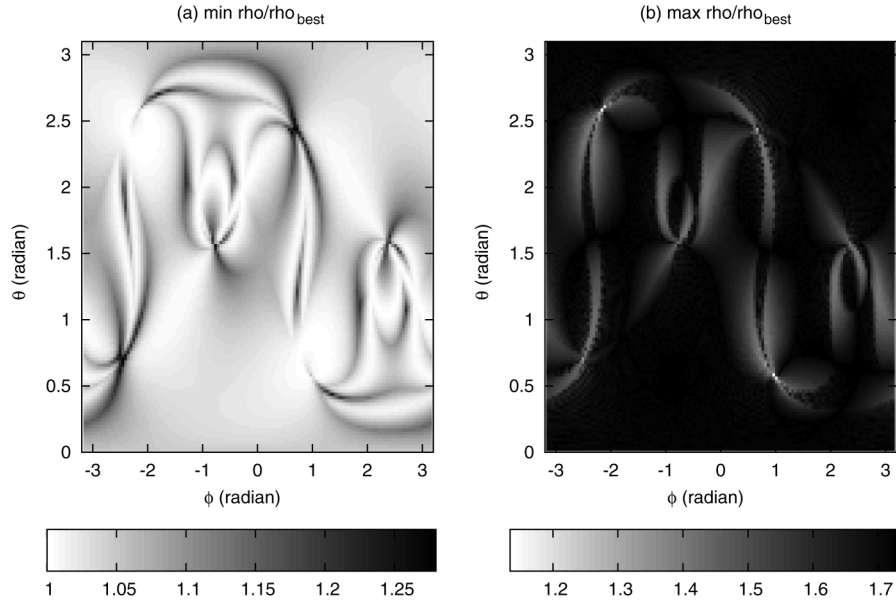


FIG. 3. *Benefits of a coherent network analysis (SNR enhancement).*—We display the polar maps of the following quantities for the L1-H1-V1 network (a) $\min_{\psi, \epsilon} \rho / \rho_{\text{best}}$ and (b) $\max_{\psi, \epsilon} \rho / \rho_{\text{best}}$. Here, ρ_{best} designates the best SNR of the detectors in the network. The maximum, minimum are taken over all the polarization angles $\{\psi, \epsilon\}$.

discussion, if we assume that we obtained parameter estimates $\hat{\mathbf{p}}$ from the data through Eq. (40), then one can retrieve the physical parameters A , ϕ_0 , ϵ , and ψ by inverting the nonlinear map which links $\mathbf{p} = (at_-^* a^* t_+ + at_+^* a^* t_-)^T$ to these parameters as given below

$$\hat{A} = (\sqrt{|\hat{\rho}[1]|} + \sqrt{|\hat{\rho}[2]|})^2, \quad (44)$$

$$\hat{\phi}_0 = \frac{1}{2}[\arg(\hat{\rho}[1]) - \arg(\hat{\rho}[2])], \quad (45)$$

$$\hat{\psi} = -\frac{1}{4}[\arg(\hat{\rho}[1]) + \arg(\hat{\rho}[2])], \quad (46)$$

$$\hat{\epsilon} = \cos^{-1} \left[\frac{\sqrt{|\hat{\rho}[2]|} - \sqrt{|\hat{\rho}[1]|}}{\sqrt{|\hat{\rho}[2]|} + \sqrt{|\hat{\rho}[1]|}} \right]. \quad (47)$$

4. Degeneracy and sensitivity of the estimate to noise

Upper bounds for the estimation error can be obtained using a perturbative analysis of the LS problem in Eq. (39). A direct use of the result of [24], Sec. 5.3.8 yields

$$\frac{\|\hat{\mathbf{p}} - \mathbf{p}\|}{\|\mathbf{p}\|} \leq \frac{\sqrt{N}}{\rho} \text{cond}(\mathbf{\Pi}). \quad (48)$$

This bound is a worst-case estimate obtained when the noise term which affects the data \mathbf{x} is essentially concentrated along the directions associated to the smallest SV of

$\mathbf{\Pi}$. The noise is random and it spans isotropically all Nd dimensions of the signal space. As described above, the space associated to the smallest SV has only 2 dimensions. Therefore, the worst case is very unlikely to occur and the above bound is largely overestimated on the average. However, it gives a general trend and shows that the estimation goes worst with the conditioning of $\mathbf{\Pi}$.

Regularization techniques seem to give promising results in the context of GW burst detection [13–15]. Following this idea, we may consider to “regularize” the LS problem in Eq. (39). To do so, additional information on the expected parameters is required to counterbalance the rank deficiency. Unfortunately, we do not expect \mathbf{p} to follow a specific structure. The only sensible prior that can be assumed without reducing the generality of the search is that $\|\mathbf{p}\|$ is likely to be bound (since the GW have a limited amplitude A). It is known [27] that this type of prior is associated to the use of the so-called Tychonov regulator and that we do not expect significant improvements upon the nonregularized solution.

One difference explains why regularization techniques do not work in the present case while it does work for burst detection. We recall that in the burst case, the parameter vector comparable to \mathbf{p} are the samples of the waveform. This vector being a time series, it is expected to have some structure, in particular, it is expected to have some degree of smoothness. The use of this *a priori* information improves significantly the final estimation.

While regularization will not help for the *estimation* of the extrinsic parameters, they may be of use to improve the *detection* statistic. We consider this separate question later in Sec. VII B.

B. Implementation with synthetic streams

In the previous section, we maximize the network LLR with respect to the extrinsic parameters resulting in the statistic $\hat{\Lambda}$ in Eq. (42). Here, we obtain a more simple and practical expression for $\hat{\Lambda}$ which will be useful for maximization over the remaining intrinsic parameters.

From Eqs. (23) and (42), we have

$$\hat{\Lambda}(\mathbf{x}) = \| (\mathbf{U}_D \otimes \mathbf{U}_E)^H \mathbf{x} \|^2. \quad (49)$$

It is useful to reshape the network data \mathbf{x} into a $N \times d$ matrix $\mathbf{X} \equiv [\mathbf{x}_1 \mathbf{x}_2 \dots \mathbf{x}_d]$. This operation is inverse to the *stack operator* $\text{vec}()$ defined in Appendix A.

Using a property of the Kronecker product in Eq. (A2), we obtain the reformulation

$$\hat{\Lambda}(\mathbf{x}) = \| \text{vec}(\mathbf{U}_E^H \mathbf{X} \mathbf{U}_D^*) \|^2. \quad (50)$$

There are two possibilities to make this matrix product, each being associated to a different numerical implementation for the evaluation of $\hat{\Lambda}$.

We can first multiply \mathbf{X} by \mathbf{U}_E^H and then by \mathbf{U}_D^* . In practice, this means that we first compute the correlation of the data with a chirp template, then we combine the result using weights (related to the antenna pattern functions). This is the implementation proposed in [11]. It is probably the best for cases (like, searches of inspiralling binary chirps) where the number of chirp templates is large (i.e., larger than the number of source locations) and where the correlations with templates are computed once and stored.

The second choice is to first multiply \mathbf{X} by \mathbf{U}_D^* and then by \mathbf{U}_E^H which we adopt here. This means that we first compute $\mathbf{Y} \equiv \mathbf{X} \mathbf{U}_D^*$ which transforms the network data into two N -dimensional complex data vectors $[\mathbf{y}_1, \mathbf{y}_2] \equiv \mathbf{Y}$ through an ‘‘instantaneous’’ linear combination. Then, we correlate these vectors with the chirp template. We can consider \mathbf{y}_1 and \mathbf{y}_2 as the output of two ‘‘virtual’’ detectors. For this reason, we refer to those as *synthetic streams* in connection to [28] who first coined the term for such combinations. Note that, irrespective of the number of detectors, one always gets at most two synthetic streams. We note that though the synthetic streams defined in [28] are *ad hoc* (i.e., they have no relation with the LR), the ones obtained here directly arise from the maximization of the network LLR.

We express the network LLR statistic in terms of the two synthetic streams as

$$\hat{\Lambda}(\mathbf{x}) = \frac{1}{N} (|\mathbf{e}^H \mathbf{y}_1|^2 + |\mathbf{e}^T \mathbf{y}_1|^2 + |\mathbf{e}^H \mathbf{y}_2|^2 + |\mathbf{e}^T \mathbf{y}_2|^2), \quad (51)$$

where $\mathbf{y}_l = \mathbf{X} \mathbf{u}_l^*$ for $l = 1, 2$. This expression can be further simplified by using the symmetry (easily seen from Eqs. (29) and (30)),

$$\mathbf{u}_1^* = \exp(i\delta) \mathbf{u}_1, \quad \mathbf{u}_2^* = -\exp(i\delta) \mathbf{u}_2. \quad (52)$$

We finally obtain

$$\hat{\Lambda}(\mathbf{x}) = \frac{2}{N} (|\mathbf{e}^H \mathbf{y}_1|^2 + |\mathbf{e}^H \mathbf{y}_2|^2). \quad (53)$$

The linear combination in each stream is such that the signal contributions from each detector add up constructively. In this sense, synthetic streams are similar to beamformers used in array signal processing [29]. The GW chirp thus appears in the synthetic streams with an enhanced amplitude. The enhancement factor can be evaluated as follows. When the data is a noise free GW chirp, i.e., $\mathbf{x} = \mathbf{s}$, we then have

$$\mathbf{y}_l[k] = \mathbf{p}^T (\mathbf{D} \otimes \mathbf{E}[k])^T \mathbf{u}_l^* = \mathbf{p}^T (\sigma_l \mathbf{v}_l^* \otimes \mathbf{E}[k]^T), \quad (54)$$

where $\mathbf{E}[k]$ represents the k th row of \mathbf{E} . Writing explicitly, we have

$$\mathbf{y}_1 = \frac{\sigma_1}{\sqrt{2}} \Re\{q_1 \tilde{\mathbf{h}}\} e^{i\delta/2}, \quad \mathbf{y}_2 = \frac{i\sigma_2}{\sqrt{2}} \Im\{q_2 \tilde{\mathbf{h}}\} e^{i\delta/2}, \quad (55)$$

where we have $q_1 = t_-^* e^{i\delta/2} + t_+^* e^{-i\delta/2}$ and $q_2 = t_-^* e^{i\delta/2} - t_+^* e^{-i\delta/2}$. This shows that the synthetic streams \mathbf{y}_l are rescaled and phase shifted copies of the initial GW chirp $\tilde{\mathbf{h}}$ as defined in Eq. (16).

SNR per synthetic streams

The network SNR can be split into the contributions from each synthetic stream i.e. using Eq. (43) we write $\rho^2 = \| \boldsymbol{\Sigma}_\Pi \mathbf{V}_\Pi^H \mathbf{p}_0 \|^2$, as

$$\rho^2 = \rho_1^2 + \rho_2^2, \quad (56)$$

where we define $\rho_l \equiv \sqrt{N} \sigma_l \| (\mathbf{v}_l \otimes \mathbf{I}_2)^H \mathbf{p}_0 \| / 2$ for $l = 1, 2$. More explicitly, we have

$$\rho_l = \frac{\sqrt{N}}{2} \sigma_l |q_l| A. \quad (57)$$

The synthetic streams contribute differently depending on the polarization of the incoming wave. Figure 4 illustrates this with the network formed by the two LIGO detectors and Virgo.

Let us assume that \mathbf{p}_0 is randomly oriented. Since \mathbf{v}_1 and \mathbf{v}_2 have unit norms, we get the average value $0 < \langle \rho_2 / \rho_1 \rangle_{\epsilon, \psi} \propto 1 / \text{cond}(\boldsymbol{\Pi}) \leq 1$ for most of the sky as indicated in Fig. 4(a). Note that this panel matches well with Fig. 2(b). Thus, on average, \mathbf{y}_1 contributes more to the SNR than \mathbf{y}_2 . However, the situation may be different depending on the specific polarization state of the wave. Figure 4(b) shows the maximum of the ratio $\rho / \max_{l=1,2} \rho_l$ for all polarization angles ϵ and ψ . For most sky locations, this quantity is $\approx \sqrt{2}$ which means that the two synthetic streams contribute equally. This holds true for all the sky locations, except at the degenerate ones where \mathbf{y}_2 does not contribute, hence the SNR ratio is 1. Inversely, one can

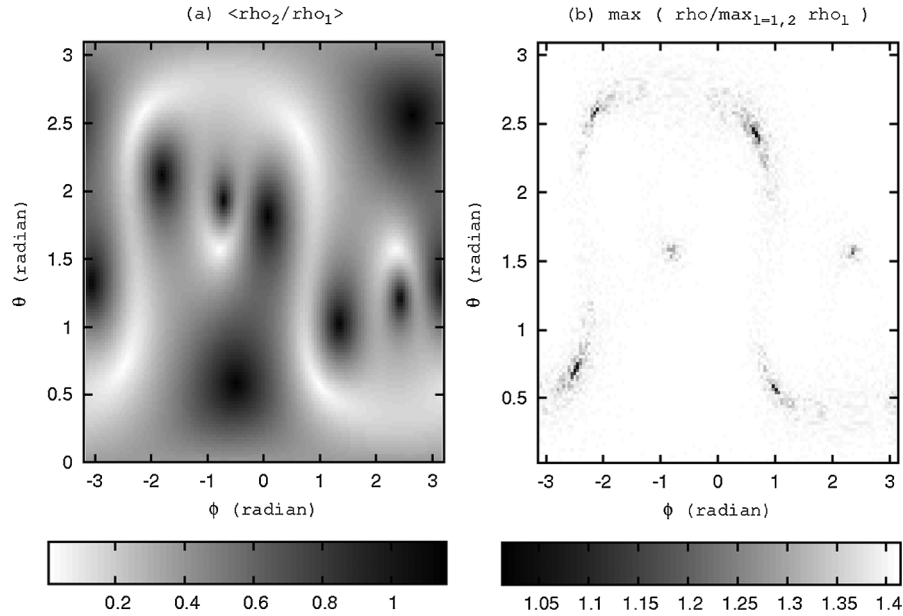


FIG. 4. *SNR per synthetic streams and benefits of a coherent network analysis (SNR enhancement).*—We display the polar maps of the following quantities for the L1-H1-V1 network: (a) $\langle \rho_2 / \rho_1 \rangle_{\psi, \epsilon}$ and (b) $\max_{\psi, \epsilon} (\rho / \max_{l=1,2} \rho_l)$ where ρ_l denotes the SNR of synthetic stream $l = 1$ or 2, as defined in Eq. (57). The maximum and average are taken over all the polarization angles $\{\psi, \epsilon\}$.

always find a GW polarization such that one of the synthetic streams does not contribute to the SNR.

Another insight into this question is given by the following expression of the difference of the SNR in each synthetic stream

$$\rho_1^2 - \rho_2^2 = \frac{NA^2}{4} \left(\frac{|\Delta|}{2} \left[\left(\frac{1 + \cos^2 \epsilon}{2} \right)^2 + \cos^2 \epsilon \right] + \frac{\|\mathbf{f}_+\|^4 - \|\mathbf{f}_\times\|^4 \sin^4 \epsilon}{2|\Delta|} \right), \quad (58)$$

where we have¹⁰ $|\Delta|^2 = (\|\mathbf{f}_+\|^2 + \|\mathbf{f}_\times\|^2)^2 - 4\|\mathbf{f}_+ \times \mathbf{f}_\times\|^2$.

For a face-on source, i.e. $\epsilon = 0$, and assuming \mathbf{f}_\times is orthogonal to \mathbf{f}_+ with $\|\mathbf{f}_\times\| = \|\mathbf{f}_+\|$, then both synthetic streams carry the same SNR $\rho_1 = \rho_2 = \sqrt{N/2A} \|\mathbf{f}_+\|$.

C. Null streams

1. Review and relation to synthetic streams

The access to *noise only* data is crucial in the signal detection problem. Such data is not directly available in GW experiments, but the use of multiple detectors allows to access it indirectly using the *null streams*. The general

¹⁰The synthetic streams (on the average sense) are also connected to the directional streams introduced in the context of LISA [30]. If we integrate ρ_l^2 over the inclination and polarization angles ϵ, ψ , we obtain $\langle (\rho_1^2 - \rho_2^2) / 2 \rangle_{\epsilon, \psi} = 2|\Delta|/5$ and $\langle \|\hat{\mathbf{f}}\|^2 \rangle_{\epsilon, \psi} = 2\mathcal{D}/5$. Thus, the SNRs of the synthetic streams ρ_l^2 when averaged over the polarization angles are proportional to the SNRs obtained by v_+ and v_\times —the directional streams in the LISA data analysis, see Eqs. (25–28) of [30].

idea behind the null stream is to construct a data stream from the individual detector streams which nullifies the signature of any incoming GW from a particular direction. Since this signal cancellation is specific to GWs, null streams naturally provide an extra tool to verify that a detected signal is indeed a GW or instead GW-like features mimicked by the detector noise whose detection thus has to be vetoed. This is a powerful check since it does not require detailed information about the potential GW signal under test, except an estimate of its source location. (Note that in practice, the implementation of the veto test may be complicated by the imprecision of the direction of arrival and of the errors of calibration [18]). The existence of null streams has been first identified in [31] in the case of three detector networks. At present, a handful of literature [17,18] exists on the use of null streams in GW data analysis.

Null streams are usually introduced as a general post-processing of the data independent of the detection of specific GWs. Below, we make this connection in the domain of our formalism. We recall that the network data at a given time (e.g., the first row of the matrix \mathbf{X} introduced in Sec. VB) is a d -dimensional vector in \mathbb{R}^d . This space is a direct sum of the GW polarization plane and its orthogonal complementary space. We have shown that the GW polarization plane is a 2-dimensional space, spanned by a pair of orthonormal basis vectors which are associated to the two synthetic streams. The complementary space to the GW polarization plane is a $d - 2$ -dimensional space and it is spanned by $d - 2$ “null vectors.” Similar to the synthetic streams, the null streams can be constructed from these null vectors. Thus, the numbers of synthetic and null

streams sum up to d . Nominally, we have $d - 2$ null streams. However, when the GW polarization plane degenerates to a 1-dimensional space ($\sigma_2 \sim 0$) as explained in Sec. IV B, the number of null vectors becomes $d - 1$. For a two detector network, in the nominal case, there is no null stream as $d - 2 = 0$. However, for degenerate directions, one can construct a null stream. For an aligned pair of detectors (as is almost the case for the two LIGO L1 and H1), the fraction of the degenerate sky location is large, see Fig. 2. This null stream would turn out to be useful for vetoing in this case.

In the next section, we explain how the null streams can be obtained numerically in the nominal case. The extension to the degenerate case is straightforward.

2. Obtaining the null streams numerically

The numerical construction of the null streams can be achieved in various ways. One such approach could be to obtain the full SVD of \mathbf{D} and construct the null streams from the eigenvectors corresponding to the zero SVs. This approach was taken in [18]. Here, we take an alternative approach. We construct the null streams by successive construction of orthonormal vectors via a multidimensional cross-product as described below.

Assuming some direction of arrival, we express any instantaneous linear combination of the time-shifted data (to compensate for different time of arrivals at the detectors' site with respect to the reference) as

$$\mathbf{y}(\mathbf{x}) \equiv \mathbf{X}\mathbf{u}, \quad (59)$$

where the vector $\mathbf{u} \in \mathbb{C}^{d \times 1}$ contains the tap coefficients.

Equation (58) can be rewritten as

$$\mathbf{y}(\mathbf{x}) = \text{vec}(\mathbf{I}_N \mathbf{X}\mathbf{u}) = (\mathbf{u}^T \otimes \mathbf{I}_N)\mathbf{x}. \quad (60)$$

The vector \mathbf{u} defines a null stream if $\mathbf{y}(\mathbf{x}) = \mathbf{0}_N$ whenever \mathbf{x} is a GW. Let us assume that we indeed observe a GW chirp i.e., $\mathbf{x} = \mathbf{s}_0 \equiv \mathbf{\Pi}\mathbf{p}_0$. We thus have

$$\mathbf{y}(\mathbf{s}_0) = [(\mathbf{u}^T \mathbf{U}_D) \otimes \mathbf{U}_E] \Sigma_{\Pi} \mathbf{V}_{\Pi}^H \mathbf{p}_0. \quad (61)$$

If \mathbf{u} is in the null space of \mathbf{U}_D , the null-stream condition is satisfied for all \mathbf{p}_0 . Since the null space of \mathbf{U}_D is orthogonal to its range, an obvious choice for \mathbf{u} is

$$\mathbf{u} = \mathbf{u}_1 \times \mathbf{u}_2 = \frac{\mathbf{d}^* \times \mathbf{d}}{\sqrt{\mathcal{D}^2 - \Delta^2}}. \quad (62)$$

Nominally, \mathbf{U}_D is a 2-dimensional plane in \mathbb{C}^d . Its null space is therefore $d - 2$ dimensional. An orthonormal basis of this space can be obtained recursively starting from $\mathbf{u}_3 = \mathbf{u}$ as defined above and applying the following generalized vector cross-product formula for $n > 3$:

$$\mathbf{u}_n[i] = \epsilon_{ijkl\dots m} \mathbf{u}_1[j] \mathbf{u}_2[k] \mathbf{u}_3[l] \dots \mathbf{u}_{n-1}[m]. \quad (63)$$

Here, $\epsilon_{ijkl\dots m}$ is the Levi-Civita symbol.¹¹ The \mathbf{u}_n denotes an orthonormal set of $d - 2$ vectors, $\{\mathbf{u}_n\}$, for $3 \leq n \leq d$. The components of these vectors are the tap coefficients to compute the null streams. By construction, the resulting null streams are uncorrelated and have the same variance.

To summarize the main features of our formalism. The representation of a GW network response of unmodeled chirp as a Kronecker product between the GW polarization plane and the chirp plane forms the main ingredient of this formalism. Such a representation allows the signal to reveal the degeneracy in a natural manner in the network response. It also evidences the two facets of the coherent network detection problem, namely, the network signal detection via synthetic streams and vetoing via null streams. The coherent formalism developed in [11] for inspiralling binaries lacked this vetoing feature due to the difference in the signal representation.

In the rest of this paper, we do not discuss/demonstrate the null streams applied as a vetoing tool to the simulated data. This will be demonstrated in the subsequent work with the real data from the ongoing GW experiments.

D. Colored noise

The formalism developed till now was exclusively targeted for the white noise case. We assumed that the noise at each detector is white Gaussian. In this subsection, we extend our formalism to the colored noise case. We remind the reader that the main focus of this paper is to develop the coherent network strategy to detect unmodeled GW chirps with an interferometric detector network. Hence, we give more emphasis on the basic formalism and keep the colored noise case with basic minimal assumption: the noise from the different detectors is colored but with the same covariance. Based on this ground work, the work is in progress to extend this to the colored noise case with different noise covariances.

Let us therefore assume now that the noise components in each detector are independent and colored, with the same covariance matrix \mathbf{R}_0 . Recall that the covariance matrix of a random vector \mathbf{a} is defined as $\mathbb{E}[(\mathbf{a} - \mathbb{E}[\mathbf{a}]) \times (\mathbf{a} - \mathbb{E}[\mathbf{a}])^H]$ where $\mathbb{E}[\cdot]$ denotes the expectation. From the independence of the noise components, the overall covariance matrix of the network noise vector is then a block-diagonal matrix, where all the blocks are identical and equal to \mathbf{R}_0 : $\mathbf{R} = \text{diag}(\mathbf{R}_0) \equiv \mathbf{I}_d \otimes \mathbf{R}_0$.

¹¹The Levi-Civita symbol is defined as

$$\begin{aligned} \epsilon_{ij\dots} &\equiv +1 \quad \text{when } i, j, \dots \text{ is an even permutation of } 1, 2, \dots, \\ &\equiv -1 \quad \text{when } i, j, \dots \text{ is an odd permutation of } 1, 2, \dots, \\ &\equiv 0 \quad \text{when any two labels are equal.} \end{aligned}$$

In this case, the network LLR becomes

$$\Lambda(\mathbf{x}) = -\|\mathbf{x} - \mathbf{\Pi}\mathbf{p}\|_{\mathbf{R}^{-1}}^2 + \|\mathbf{x}\|_{\mathbf{R}^{-1}}^2, \quad (64)$$

where the notation $\|\cdot\|_{\mathbf{R}^{-1}}$ denotes the norm induced by the inner product associated to the covariance matrix \mathbf{R}^{-1} , i.e., $\|\mathbf{a}\|_{\mathbf{R}^{-1}}^2 = \mathbf{a}^H \mathbf{R}^{-1} \mathbf{a}$.

Introducing the whitened version $\tilde{\mathbf{\Pi}} = \mathbf{R}^{-1/2} \mathbf{\Pi}$ and $\tilde{\mathbf{x}} = \mathbf{R}^{-1/2} \mathbf{x}$ of $\mathbf{\Pi}$ and \mathbf{x} , respectively, Eq. (65) can be rewritten as

$$\Lambda(\mathbf{x}) = -\|\tilde{\mathbf{x}} - \tilde{\mathbf{\Pi}}\mathbf{p}\|^2 + \|\tilde{\mathbf{x}}\|^2, \quad (65)$$

which is similar to Eq. (38) where all the quantities are replaced by their whitened version. Thus, the maximization of $\Lambda(\mathbf{x})$ with respect to the extrinsic parameters \mathbf{p} will follow the same algebra as that derived in V. However, for the sake of completeness, we detail it below.

Following Sec. V, maximizing $\Lambda(\mathbf{x})$ with respect to the extrinsic parameters \mathbf{p} leads to

$$\hat{\mathbf{p}} = \tilde{\mathbf{\Pi}}^\# \tilde{\mathbf{x}}, \quad (66)$$

where $\tilde{\mathbf{\Pi}}^\#$ is the pseudoinverse of $\tilde{\mathbf{\Pi}}$. Expressing this pseudoinverse by means of the SVD of $\tilde{\mathbf{\Pi}}$ as $\tilde{\mathbf{\Pi}}^\# = \mathbf{V}_{\tilde{\mathbf{\Pi}}} \Sigma_{\tilde{\mathbf{\Pi}}}^{-1} \mathbf{U}_{\tilde{\mathbf{\Pi}}}^H$ and introducing Eq. (67) into Eq. (66) provides the new statistic

$$\hat{\Lambda}(\mathbf{x}) = \|\mathbf{U}_{\tilde{\mathbf{\Pi}}}^H \tilde{\mathbf{x}}\|^2. \quad (67)$$

Now, from the definition of $\mathbf{\Pi}$ and the specific structure of \mathbf{R} , it is straightforward to see that

$$\tilde{\mathbf{\Pi}} = \mathbf{R}^{-1/2} (\mathbf{D} \otimes \mathbf{E}) = \mathbf{D} \otimes (\mathbf{R}_0^{-1/2} \mathbf{E}) = \mathbf{D} \otimes \tilde{\mathbf{E}}, \quad (68)$$

where we have introduced the whitened version $\tilde{\mathbf{e}} = \mathbf{R}_0^{-1/2} \mathbf{e}$ of the chirp signal and the corresponding matrix $\tilde{\mathbf{E}} = [\tilde{\mathbf{e}} \tilde{\mathbf{e}}^*] / 2$.

For the white noise case, the statistic (68) can then be rewritten in terms of the SVD of the matrices \mathbf{D} and $\tilde{\mathbf{E}}$:

$$\hat{\Lambda}(\mathbf{x}) = \|(\mathbf{U}_D \otimes \mathbf{U}_{\tilde{\mathbf{E}}})^H \tilde{\mathbf{x}}\|^2. \quad (69)$$

The computation of $\mathbf{U}_{\tilde{\mathbf{E}}}$ is similar to the computation of $\mathbf{U}_{\tilde{\mathbf{e}}}$. Furthermore, if we note that $\tilde{\mathbf{e}}^T \tilde{\mathbf{e}} \approx 0$, and if we assume that $\tilde{\mathbf{e}}^H \tilde{\mathbf{e}} = N$, it turns out that $\mathbf{U}_{\tilde{\mathbf{E}}} = 2\tilde{\mathbf{E}} / \sqrt{N} = 2\mathbf{R}_0^{-1/2} \mathbf{E} / \sqrt{N}$. Using the property of the Kronecker product in Eq. (A2), we then obtain

$$\hat{\Lambda}(\mathbf{x}) = \|\text{vec}(\mathbf{E}^H \tilde{\mathbf{X}} \mathbf{U}_D^*)\|^2, \quad (70)$$

where the matrix $\tilde{\mathbf{X}} = [\tilde{\mathbf{x}}_1 \tilde{\mathbf{x}}_2 \dots \tilde{\mathbf{x}}_d]$ contains the data vector from each detector whitened twice: $\tilde{\mathbf{x}}_j = \mathbf{R}^{-1/2} \mathbf{x}_j = \mathbf{R}^{-1} \mathbf{x}_j$.

As this expression is similar to the white noise case, we can form two synthetic streams $[\tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2] = \tilde{\mathbf{X}} \mathbf{U}_D^*$ and use them to express the LLR statistic as

$$\hat{\Lambda}(\mathbf{x}) = \frac{2}{N} (|\mathbf{e}^H \tilde{\mathbf{y}}_1|^2 + |\mathbf{e}^H \tilde{\mathbf{y}}_2|^2). \quad (71)$$

In this expression, the only difference with the white noise LLR of (53) comes from the computation of the synthetic streams $\tilde{\mathbf{y}}_1$ and $\tilde{\mathbf{y}}_2$ which are obtained after double-whitening the data.

VI. MAXIMIZATION OVER THE INTRINSIC PARAMETERS

In the previous section, we maximized the network LLR over the extrinsic parameters of the signal model, assuming that the remaining parameters (the source location angles ϕ and θ and the phase function $\varphi(\cdot)$) were known.

By definition, the intrinsic parameters modify the network LLR nonlinearly. For this reason, the maximization of $\hat{\Lambda}$ over these parameters is more difficult. It cannot be done analytically and must be performed numerically, for instance with an exhaustive search of the maximum by repeatedly computing $\hat{\Lambda}$ over the entire range of possibilities.

While the exhaustive search can be employed for the source location angles, it is not applicable to the chirp phase function, which requires a specific method. For the single detector case, we had addressed this issue in [8] with an original maximization scheme which is the cornerstone of the BCC algorithm. Here, we use and adapt the principles of BCC to the multiple detector case.

A. Chirp phase function

Let us examine first the case of the detection of inspiraling binary chirps. In this case, the chirp phase is a prescribed function of a small number of parameters i.e., the masses and spins of the binary stars. The maximization over those is performed by constructing a grid of reference or *template* waveforms which are used to search the data. This grid samples the range of the physical parameters. This sampling must be accurate (the template grid must be tight) to avoid missing any chirp.

A tight grid of templates can be obtained in the non-parametric case (large number of parameters) i.e., when the chirp is not completely known. We have shown in [8] how to construct a template grid which covers entirely the set of smooth chirps i.e., chirps whose frequency evolution has some regularity as described in Sec. IIB 2. In the next section, we briefly describe this construction.

1. Chirplet chains: tight template bank for smooth chirps

We refer to the template forming this grid as *chirplet chain* (CC). These CCs are constructed on a simple geometrical idea: a broken line is a good approximation of a smooth curve. Since the frequency of a smooth chirp follows a smooth frequency vs time curve, we construct

templates that are broken lines in the time-frequency (TF) plane.

More precisely, CC are defined as follows. We start by sampling the TF plane with a regular grid consisting of N_t time bins and N_f frequency bins. We build the template waveforms like a puzzle by assembling small chirp pieces which we refer to as *chirplets*. A chirplet is a signal with a frequency joining linearly two neighboring vertices of the grid. The result of this assembly is a *chirplet chain* i.e., a piecewise linear chirp. Since we are concerned with continuous frequency evolution with bounded variations, we only form continuous chains.

We control the variations of the CC frequency. The frequency of a single chirplet does not increase or decrease more than N'_f frequency bins over a time bin. Similarly, the difference of the frequency variations of two successive chirplets in a chain does not increase or decrease more than N''_f frequency bins.

The CC grid is defined by four parameters namely N_t , N_f , N'_f , and N''_f . Those are the available degrees of freedom we can tune to make the CC grid tight. A template grid is tight if the network ambiguity $\hat{\Lambda}(\mathbf{s}; \varphi') \propto |\mathbf{e}^{iH}\mathbf{y}_1(\mathbf{s})|^2 + |\mathbf{e}^{iH}\mathbf{y}_2(\mathbf{s})|^2$ which measures the similarity between an arbitrary chirp (of phase φ) and its closest template (of phase φ'), is large enough and relatively closed to the maximum (when $\varphi' = \varphi$).

As stated in Eq. (55), in the presence of a noise free GW chirp, the synthetic streams $\mathbf{y}_i(\mathbf{s})$ are rescaled and phase shifted copies of the initial GW chirp \mathbf{s} . Therefore, we treat the network ambiguity as a sum of ambiguities from two *virtual* detectors where each term in $\hat{\Lambda}(\mathbf{s}; \varphi')$ is the ambiguity computed in the single detector case. An estimate of the ambiguity has been obtained in [8] for this case. It can thus be directly reused to compute $\hat{\Lambda}(\mathbf{s}; \varphi')$.

The bottom line is that the ratio of the ambiguity to its maximum for the network case remains unaltered as compared to the single detector case and thus the same for the tight grid conditions. In conclusion, the rules (which we will not repeat here) established in [8] to set the search parameters can also be applied here.

2. Search through CCs in the time-frequency plane: best network CC algorithm

We have now to search through the CC grid to find the best matching template, i.e., which maximizes $\hat{\Lambda}(\mathbf{x}; \hat{\varphi}')$ over all CCs of phase φ' . Counting the number of possible CCs to be searched over is a combinatorial problem. This count grows exponentially with the number of time bins N_t . In the situation of interest, it reaches prohibitively large values. The family of CCs cannot be scanned exhaustively and the template based search is intractable.

In [8], we propose an alternative scheme yielding a close approximation of the maximum for the single detector statistic. When applied to the network, the scheme de-

mands to reformulate the network statistic in the TF plane. The TF plane offers a natural and geometrically simple representation of chirp signals which simplifies the statistic. It turns out that the resulting statistic falls in a class of objective functions where efficient combinatorial optimization algorithms can be used. We now explain this result in more detail.

We use the TF representation given by the discrete Wigner-Ville (WV) distribution [32] defined for the time series $x[n]$ with $n = 0, \dots, N - 1$ as

$$w_x(n, m) \equiv \sum_{k=-k_n}^{+k_n} x[[n + k/2]]x^*[[n - k/2]]e^{-2\pi imk/(2N)}, \quad (72)$$

with $k_n \equiv \min\{2n, 2N - 1 - 2n\}$, where $[\cdot]$ gives the integer part. The arguments of w_x are the time index n and the frequency index m which correspond, in physical units, to the time $t_n = t_s n$ and the frequency $\nu_m = \nu_s m/(2N)$ for $0 \leq m \leq N$ and $\nu_m = \nu_s(N - m)/(2N)$ for $N + 1 \leq m \leq 2N - 1$.

The above WV distribution is a unitary representation. This means that the scalar products of two signals can be reexpressed as scalar products of their WV. Let $x_1[n]$ and $x_2[n]$ be two time series. The unitarity property of w_x is expressed by the Moyal's formula as stated below

$$\left| \sum_{n=0}^{N-1} x_1[n]x_2^*[n] \right|^2 = \frac{1}{2N} \sum_{n=0}^{N-1} \sum_{m=0}^{2N-1} w_{x_1}(n, m)w_{x_2}(n, m). \quad (73)$$

Applying this property to the network statistic in Eq. (53), we get

$$\hat{\Lambda}(\mathbf{x}) = \frac{1}{N^2} \sum_{n=0}^{N-1} \sum_{m=0}^{2N-1} w_y(n, m)w_e(n, m), \quad (74)$$

where $w_y = w_{y_1} + w_{y_2}$ combines the individual WVs of the two synthetic streams.

In order to compute $\hat{\Lambda}(\mathbf{x})$, we need to have a model for w_e . We know that the WV distribution of a linear chirp (whose frequency is a linear function of time) is essentially concentrated in the neighborhood of its instantaneous frequency [32]. We assume that it also holds true for an arbitrary (nonlinear) chirp. Applying this approximation to the WV w_e of the template CC in Eq. (75), we get

$$w_e(n, m) \approx 2N\delta(m - m_n), \quad (75)$$

where m_n denotes the nearest integer of $2T\nu(t_n)$ and ν is the instantaneous frequency of the CC.

Thus, substituting in Eq. (75), we obtain the following reformulation of the network statistic

$$\hat{\Lambda}(\mathbf{x}) = \frac{2}{N} \sum_{n=0}^{N-1} w_y(n, m_n). \quad (76)$$

The maximization of $\hat{\Lambda}(\mathbf{x})$ over the set of CC amounts to finding the TF path that maximizes the integral Eq. (77), which is equivalent to a *longest path problem* in the TF plane. This problem is structurally identical to the single detector case (the only change is the way we obtain the TF map). We can therefore essentially reuse the scheme proposed earlier for this latter case. The latter belongs to a class of combinatorial optimization problems where efficient (polynomial time) algorithms exist. We use one such algorithm, namely, the dynamic programming.

In conclusion, *the combination of the two ingredients, namely, the synthetic streams and the phase maximization scheme used in BCC allows us to coherently search the unmodeled GW chirps in the data of the GW detector network.* We refer to this procedure as the BNCC algorithm.

B. Source sky position

As we are performing maximization successively, till now we assume that we know the sky position of the source. Knowing the sky position, we construct the synthetic streams with appropriate direction dependent weight factors, time-delay shifts, and carry out the BNCC algorithm for chirp phase detection. In reality, the sky position is unknown. One needs to search through the entire sky by sampling the celestial sphere with a grid and repeating the above procedure for each point on this grid.

C. Time of arrival

Since we process the data streams sequentially and blockwise, the maximization over t_0 amounts to selecting that block where the statistic arrives at a local maximum (i.e., the maximum of the “detection peak”). The epoch of this block yields an estimate of t_0 . The resolution of the estimate may be improved by increasing the overlap between two consecutive blocks.

D. Estimation of computational cost

We estimate the computational cost of the BNCC search by counting the floating-point operations (flops) required by its various subparts. The algorithm consists essentially in repeating the one-detector search for all sky location angles. Let N_Ω be the number of bins of the sky grid. The total cost is therefore N_Ω times the cost of the one-detector search, which we give in [8] and summarize now. The computation of the WV of the two synthetic streams requires $10N_f \log_2 N_f$ flops and the BCC search applied to the combined WV requires $[N + (2N_r'' + 1)N_r']N_c$ flops, where $N_c \approx (2N_r' + 1)N_f$ is the total number of chirplets. Since this last part of the algorithm dominates, the overall cost thus scales with

$$C \propto N_\Omega [N + (2N_r'' + 1)N_r'] (2N_r' + 1)N_f. \quad (77)$$

This is the numerical cost for computing one data block of duration T . The computational power needed to process the data in real time is thus given by $C/(\mu T)$ where μ is the overlap between two successive blocks.

VII. RESULTS WITH SIMULATED DATA AND DISCUSSION

A. Proof of principle of a full blind search

We present here a proof of principle for the proposed detection method. For this case study, we consider a network of three detectors placed and oriented like the existing Virgo and the two four-kilometer LIGO detectors. The coordinates and orientation of these detectors can be found in Table I. We assume a simplified model for detector noise which we generate independently for each detector, using a white Gaussian noise. Figure 5 illustrates the possibility of a “full blind” search in this situation. This means that we perform the detection jointly with the estimation of the GW chirp frequency and the source sky location.

1. Description of the test signal

Because of computational limitation, we restrict this study to rather short chirps of $N = 256$ samples, i.e., a chirp duration $T = 250$ ms assuming a sampling rate of $\nu_s = 1024$ Hz. The chirp frequency follows a random time evolution which, however, satisfies chirping rate constraints. We make sure that the first and second derivatives of the chirp frequency are not larger than $F' = 9.2$ kHz/s and $F'' = 1.57$ MHz/s², respectively. The chosen test signal has about 50 cycles. This is a larger number than what is considered typically for burst GWs (~ 10).

As a comparison with a well-known physical case, an inspiralling (equal mass) binary with total mass $M = 11M_\odot$ reaches the same maximum frequency variations at the last stable circular orbit. (Binary chirps with larger total mass also satisfy these chirping rate limits).

We set the SNR to $\rho = 20$. The chirp is injected at the sky position $\phi = 2.8$ rad and $\theta = 0.4$ rad where the contributions of the individual detectors are comparable, namely, the individual SNR are 10.4, 10.15, and 13.77 for Virgo, LIGO Hanford, and LIGO Livingstone, respectively.

2. Search parameters

We search through the set of CCs defined over a TF grid with $N_t = 128$ time intervals and $N = N_f = 256$ frequency bins (using $f_s = 2048$ Hz). We set the regularity parameters to $N_r' = 9$ and $N_r'' = 3$, consistent to the above chirping rate limits.

We select an *ad hoc* sky grid by dividing regularly the full range of the source localization angles θ and ϕ into 128 bins. The resulting grid has therefore a total of $N_\Omega = 16\,384$ bins. This is probably much finer than is required to perform the detection without missing candidate. However,

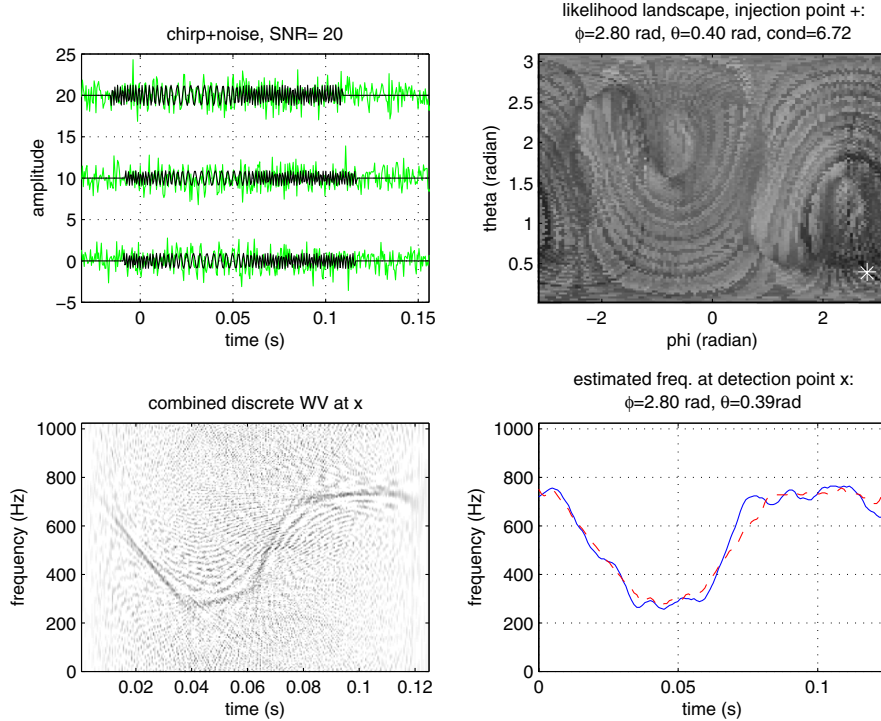


FIG. 5 (color online). *Coherent detection/estimation of a “random” chirp with network of three GW antennas.*—In the data of three GW antennas (the two LIGO and Virgo), we inject (a) a random GW chirp emitted from a source at the position marked with a “+” at $\phi = 2.8$ rad and $\theta = 0.4$ rad. We perform a full-sky search using the best network chirplet chain algorithm. It produces a likelihood landscape (b) where we select the maximum. This is the detection point and it is indicated with “×”. In (c) we show the combined WV distribution of the synthetic streams at the detection point. In (d), we compare the exact frequency of the chirp (solid/blue line) with the estimation (dashed/red line) obtained at the detection point.

this oversampling leads to precise likelihood sky maps which helps to diagnose the method. With this parameter choice, the estimated computational power required to analyze the data in real time is of 2.8 TFlops, assuming an overlap of $\mu = 50\%$ between successive data blocks. Because of the crude choice for the sky grid, this requirement is probably overestimated.

The result of the search is displayed in Fig. 5 where we see that the injection is recovered both in sky position and frequency evolution. The source position is estimated at $\hat{\phi} = 2.8$ rad and $\hat{\theta} = 0.39$ rad.

B. Regularized variants

As shown in Sec. , the SNR carried by the synthetic stream is proportional to the corresponding SV. When the GW polarization plane is degenerate (i.e., when σ_2 is small), the second synthetic stream contains almost only noise. We thus do not lose information if we suppress its contribution from the statistic. This is the basic idea of Klimenko *et al.* in [13].

We have seen that the estimation of the extrinsic parameters is an ill-posed least-square problem in those cases. Suppressing the contribution of the second synthetic

stream amounts to *regularizing* this problem [13]. In practice, this regularization can be done in various ways, corresponding to well-identified schemes.

A first possibility is to suppress the contribution of the second synthetic stream when the conditioning number of $\mathbf{\Pi}$ is too large (i.e., exceeds a given threshold). This scheme is referred to as *truncated SVD* [27]. A second possibility is to balance (divide) the contribution of the second synthetic stream by the conditioning number. This is referred to as the *Tykhonov* approach [27] and it was proposed for regularizing burst searches in [15].

In Fig. 6, we compare the likelihood landscape and frequency estimate obtained with the standard statistic and its regularized version using the Tykhonov approach. Visually, the regularization improves the contrast and concentration of the likelihood landscape around the injection point. This can be assessed more quantitatively with the contrast defined as the ratio of the likelihood landscape extremes. This contrast is improved by about 10% for the regularized statistic as compared to the standard version. It is also interesting to compare the “width” of the detection peak obtained with the two statistic. To do this, we measure the solid angle of the sky region where the statistic is larger than 90% of the maximum. This angle is reduced by a

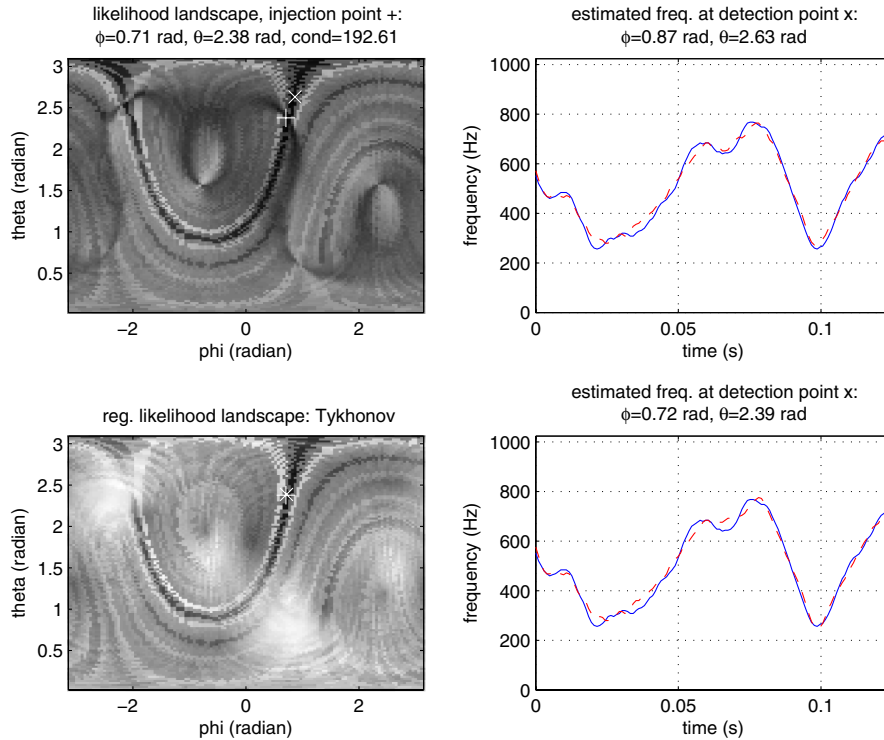


FIG. 6 (color online). *Coherent detection/estimation of a random chirp with network of three GW antennas. Standard and regularized statistic.*—We compare the likelihood landscape (left) and frequency estimation (right) obtained using the standard (top) and Tykhonov-regularized (bottom) versions of the network statistic. The test signal is a random GW chirp injected at the sky location marked with a “+”. This location has been chosen because of the associated large value of the conditioning number, namely $\text{cond}(\mathbf{I}) \approx 192$.

factor of ~ 6 when computed with the regularized statistic. There is, however, no major improvement of the frequency estimate. More generally, it is unclear whether the regularized statistic performs better than the standard one.

VIII. CONCLUDING REMARKS

The coherent detection of unmodeled chirps with a network of GW have features and issues in common with the burst one. In particular, the same geometrical objects play a key role. While the noise spans the whole d -dimensional data space, GW signals (chirps or burst) only belong to a two-dimensional (one dimension per GW polarizations) subspace, the *GW polarization plane*. Detecting GWs amounts to checking whether the data has significant components in this plane or not. To do so, we compute the projections of the data onto a basis of the GW polarization plane. In practice, this defines two instantaneous linear mixtures of the individual detector data which we refer to as *synthetic streams*. Those may be considered as the output data of virtual detectors. This combination is such that the GW contributions from each real detector add constructively. The GW signature thus has a larger amplitude

in the synthetic streams while the noise variance is kept at the same level.

The coherent detection amounts to looking for an excess in the signal energy in one or both synthetic streams (depending on the GW polarization model). This provides a generic and simple procedure to produce a coherent detection pipeline from a one-detector pipeline. In the one-detector case, the BCC search performs a path search in a time-frequency distribution of the data. In the multiple detector case, the BNCC search now uses the joint time-frequency map obtained by summing the time-frequency energy distributions of the two synthetic streams. The approach does not restrict to chirp detection and it can be applied to burst searches [18].

We demonstrated in a simplified situation that the full-sky blind detection of an unmodeled chirp is feasible. This means that the detection is performed jointly to estimate the source location and the frequency evolution. The application of this method to the real data, however, requires several improvements. First, the method has to be adapted to the case where the detectors have different sensitivities. In this respect, we already obtained first results [33].

We also have to refine the choice of the grid which samples the celestial sphere. In the present work, we select

an *ad hoc* grid. Clearly, the sky resolution and the bin shape depends on the geometry of the sphere and the location and orientation of the detectors in the considered network. A better grid choice (not too coarse to avoid SNR losses, and not too fine to avoid using useless computing resources) should incorporate this information keeping the search performance (detection probability and sky resolution) constant. In this respect, we may consider to other parametrizations of the sky location which makes the definition of the sky grid easier, for instance by choosing the time delays as investigated in [11]. We may also explore hierarchical schemes for the reduction of the computational cost.

The GW polarization plane depends on the detector antenna patterns functions. With the presently available networks, there are significantly large sky regions where the antenna patterns are almost collinear. In this case, the network observes essentially one polarization and is almost insensitive to the other: the GW polarization plane reduces to a one-dimensional space. The information carried by the missing polarization lacks and this makes the estimation of certain parameters ill-posed and hence very sensitive to noise. We can evaluate that the variance of the estimate scales with the condition number of the antenna pattern matrix. When this number (which quantifies in some sense the mutual alignment of the detectors) is large, the estimation is ill-posed and we expect poor results.

This is an important issue for burst detection since it affects significantly the shape of the estimated waveform (and, particularly the regularity of its time evolution). This has motivated the development of regularization schemes which penalize the estimation of nonphysical (i.e., irregular) waveforms. We have shown that this is, however, less of a problem for chirps because of their more constrained model. Ill-conditioning only affects global scaling factors in the chirp model. Unlike bursts, no additional prior is available for regularizing the estimation of these scaling factors.

The data space can be decomposed as the direct sum of the GW polarization plane and its complementary. While GWs have zero components in the latter null space, it is unlikely that instrumental noise (including its non-Gaussian and nonstationary part) will. This motivates the use of *null streams* (i.e., the projection of the data along a basis of the complementary space) to verify that a trigger is indeed a GW candidate and not an instrumental artifact. Since null streams are inexpensive to compute, we consider to use them to make preemptive cuts in order to avoid the analysis of bad data.

ACKNOWLEDGMENTS

A.P. is supported by the Alexander von Humboldt Foundation's Sofja Kovalevskaja Programme (funded by the German Ministry of Education and Research). O.R. is supported by the Virgo-EGO Scientific Forum. The authors would like to thank J.-F. Cardoso for an interesting exchange of ideas.

APPENDIX A: KRONECKER PRODUCT: DEFINITION AND PROPERTIES

The Kronecker product \otimes transforms two matrices $\mathbf{A} \in \mathbb{C}^{m \times n}$ and $\mathbf{B} \in \mathbb{C}^{p \times q}$ into the following matrix of $\mathbb{C}^{mp \times nq}$ [25]

$$\mathbf{A} \otimes \mathbf{B} \equiv \begin{bmatrix} a_{11}\mathbf{B} & \dots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{m1}\mathbf{B} & \dots & a_{mn}\mathbf{B} \end{bmatrix}. \quad (\text{A1})$$

The Kronecker product is a linear transform and can be considered as a special case of the tensor product. We define the operator $\text{vec}()$ to be the stack operator which transforms the matrix into a vector by stacking its columns, i.e. $\mathbf{x} \equiv \text{vec}(\mathbf{X})$. In the text, we use the following property:

$$(\mathbf{A} \otimes \mathbf{B})\text{vec}(\mathbf{X}) = \text{vec}(\mathbf{B}\mathbf{X}\mathbf{A}^T). \quad (\text{A2})$$

The proof of this property is straightforward.

APPENDIX B: INTERFEROMETRIC DETECTOR RESPONSE IN TERMS OF GEL'FAND FUNCTIONS

The GW response of a detector to an incoming GW can be obtained by computing the interaction of the wave tensor W with the detector tensor D as follows¹²:

$$s = \sum_{i,j=1}^3 W^{ij} D_{ij}. \quad (\text{B1})$$

The wave tensor is related to the incoming GW tensor in the TT gauge by $h_{ij} = 2W_{ij}$. Both detector and wave tensors are rank 2 STF tensors. Any STF tensor can be expanded in the basis of spin-weighted spherical harmonics of rank 2 and the rank-2 Gel'fand functions provide the corresponding coefficients. Further, they are representation of rotation group SO(3) and provide compact representation for the detector response of any arbitrarily oriented and located detector on Earth which we present in this appendix [34].

a. Wave tensor—The incoming GW tensor in TT gauge is given by

$$h_{ij}^{TT} = (e_{xi}e_{xj} - e_{yi}e_{yj})h_+ + 2(e_{xi}e_{yj})h_\times, \quad (\text{B2})$$

where \mathbf{e}_x and \mathbf{e}_y are unit vectors along the x_w and y_w axes in the wave frame; h_+ and h_\times are the two GW polarizations. Let $\hat{m} = (\mathbf{e}_x + i\mathbf{e}_y)/\sqrt{2}$ be a complex vector in the wave frame. Then, the wave tensor can be written down in terms of m as

$$W^{ij} = \Re[(m^i m^j)^* h], \quad (\text{B3})$$

where we used the complex quantity $h = h_+ + ih_\times$ which combines both GW polarizations.

¹²Unless otherwise mentioned, the notations and symbols used in all appendices are confined to those appendices only.

The term $m_i m_j$ is a STF tensor of rank 2. We choose to work in the detector frame for convenience. Expanding $m_i m_j$ in terms of spin-weighted spherical harmonics of rank 2, namely \mathcal{Y}_{2n}^{ij} and the rank-2 Gel'fand functions $T_{mn}(\phi'_e, \theta'_e, \psi'_e)$, we get

$$m^i m^j = \sqrt{\frac{8\pi}{15}} \sum_{n=-2}^2 \mathcal{Y}_{2n}^{ij} T_{2n}(\phi'_e, \theta'_e, \psi'_e). \quad (\text{B4})$$

The angles $(\phi'_e, \theta'_e, \psi'_e)$ are the Euler angles of the rotation operator which transforms the detector coordinates into the wave coordinates.

Substituting Eq. (B4) into Eq. (B3), we express the wave tensor in terms of the Gel'fand functions as

$$W^{ij} = \sqrt{\frac{8\pi}{15}} \sum_{n=-2}^2 \Re[h \mathcal{Y}_{2n}^{ij} T_{-2n}]. \quad (\text{B5})$$

b. Detector tensor—The detector tensor is

$$D_{ij} = n_{1i} n_{1j} - n_{2i} n_{2j}, \quad (\text{B6})$$

where \mathbf{n}_1 and \mathbf{n}_2 are the unit vectors along the first and second arms of the interferometer. Recall that we choose the x_d -axis of the detector frame along the bisector of the two arms. The y_d -axis is chosen such that (x_d, y_d, z_d) is a right-handed coordinate system with z_d pointing towards the local zenith. With this choice, we have

$$D_{11} = D_{22} = 0, \quad D_{12} = D_{21} = -1. \quad (\text{B7})$$

From Eqs. (B6) and (B7), the detector response is

$$s = \Re[f^* h], \quad (\text{B8})$$

where the complex antenna pattern function is given by

$$f = i[T_{2-2}(\phi'_e, \theta'_e, \psi'_e) - T_{22}(\phi'_e, \theta'_e, \psi'_e)]. \quad (\text{B9})$$

From the expansion of the above Eq. (B8) in terms of GW polarizations, it is consistent to define $f = f_+ + i f_\times$, which yields Eq. (13).

c. Extended complex antenna pattern for sources orbiting in a fixed plane—As discussed in Sec. III D 1, the extended antenna pattern functions incorporate the inclination angle ϵ and is given in Eq. (17) as

$$\tilde{f} = \frac{1 + \cos^2 \epsilon}{2} f_+ + i \cos \epsilon f_\times, \quad (\text{B10})$$

where $f_{+,\times}$ depend on the relative orientation of wave frame with respect to the detector frame. The detector to wave frame coordinate transformation can be split into two: detector to Earth's frame and Earth's frame to wave frame by the following rotation transformations as given in Eqs. (5) and (6)

$$\mathcal{O}(\phi'_e, \theta'_e, \psi'_e) = \mathcal{O}(\phi_e, \theta_e, \psi_e) \mathcal{O}^{-1}(\alpha, \beta, \gamma). \quad (\text{B11})$$

The above successive rotation transformation can be translated into the addition theorem of Gel'fand functions [35]

as given below

$$T_{mn}(\phi'_e, \theta'_e, \psi'_e) = \sum_{l=-2}^2 T_{ml}(\phi_e, \theta_e, \psi_e) T_{nl}^*(\alpha, \beta, \gamma). \quad (\text{B12})$$

We used the fact that the inverse rotation operator is associated to a complex conjugation.

Substituting in Eq. (B9), we rewrite the antenna pattern functions in terms of the Gel'fand functions as

$$f = - \sum_{s=-2}^2 iT_{2s}(\phi_e, \theta_e, \psi_e) \times [T_{2s}(\alpha, \beta, \gamma) - T_{-2s}(\alpha, \beta, \gamma)]^*. \quad (\text{B13})$$

Substituting in the extended beam pattern function given in Eq. (B10) and combining the dependencies upon $\psi = \psi_e$ and ϵ , we get

$$\tilde{f} = T_{22}(\psi, \epsilon, 0)d + T_{2-2}(\psi, \epsilon, 0)d^*, \quad (\text{B14})$$

where

$$d = - \sum_{n=-2}^2 iT_{2n}(\phi_e, \theta_e, 0)[T_{2n}(\alpha, \beta, \gamma) - T_{-2n}(\alpha, \beta, \gamma)]^*. \quad (\text{B15})$$

There are various ways of expressing the antenna pattern functions. The main advantages of this one is that it is particularly compact and that the angles ψ and ϵ get factorized from the rest of the parameters. This helps in the maximization of the network LLR over the extrinsic parameters.

APPENDIX C: SVD OF A TWO-COLUMN COMPLEX MATRIX

In this appendix, we obtain the SVD of a complex matrix of the type $\mathbf{A} \equiv [\mathbf{a}, \mathbf{a}^*]$ where $\mathbf{a} \in \mathbb{C}^{N \times 1}$. The SVD decomposes \mathbf{A} into the product $\mathbf{A} = \mathbf{U}_A \mathbf{\Sigma}_A \mathbf{V}_A^H$ where \mathbf{U}_A and \mathbf{V}_A are two orthogonal matrices and $\mathbf{\Sigma}_A$ is a positive definite diagonal matrix. To obtain it analytically, we first get the eigen-decomposition of

$$\mathbf{A}^H \mathbf{A} = \begin{bmatrix} \mathbf{a}^H \mathbf{a} & \mathbf{a}^H \mathbf{a}^* \\ \mathbf{a}^T \mathbf{a} & \mathbf{a}^T \mathbf{a}^* \end{bmatrix} \equiv \begin{bmatrix} a & b^* \\ b & a \end{bmatrix}. \quad (\text{C1})$$

We distinguish two cases depending on the value of b .

1. For $|b| > 0$

The eigenvalues σ_\pm and the eigenvectors \mathbf{v}_\pm of $\mathbf{A}^H \mathbf{A}$ are given below:

$$\sigma_\pm^2 = a \pm |b|, \quad (\text{C2})$$

$$\mathbf{v}_\pm = \frac{1}{\sqrt{2}} \begin{bmatrix} \exp(-i\varpi) \\ \pm 1 \end{bmatrix}, \quad (\text{C3})$$

where $\varpi = \arg b = \arg[\mathbf{a}^T \mathbf{a}]$. The number of nonzero eigenvalues of $\mathbf{A}^H \mathbf{A}$ gives the rank of \mathbf{A} . The eigenvalues arranged in the descending order form the diagonal values of Σ_A^2 as shown below. The eigenvectors \mathbf{v}_\pm are the right-handed singular vectors of \mathbf{A} and form \mathbf{V}_A as given below:

$$\Sigma_A = \begin{bmatrix} \sigma_+ & 0 \\ 0 & \sigma_- \end{bmatrix}, \quad (\text{C4})$$

$$\mathbf{V}_A = [\mathbf{v}_+ \quad \mathbf{v}_-]. \quad (\text{C5})$$

We then form the matrix $\mathbf{U}_A = \mathbf{A} \mathbf{V}_A \Sigma_A^{-1} \equiv [\mathbf{u}_+, \mathbf{u}_-]$ containing the right-handed singular vectors, with

$$\mathbf{u}_\pm = \frac{\exp(-i\varpi) \mathbf{a} \pm \mathbf{a}^*}{\sqrt{2}\sigma_\pm}. \quad (\text{C6})$$

The above expressions are valid only when the two SV are nonzero. It is possible that the smallest SV σ_- vanishes. In this case, the SVD collapses to $\mathbf{U}_A = \mathbf{u}_+$, $\Sigma_A = \sigma_+$, and $\mathbf{V}_A = \mathbf{v}_+$.

2. For $|b| = 0$

In this case, the matrix $\mathbf{A}^H \mathbf{A} = a \mathbf{I}_2$ is diagonal. We thus have $\Sigma = \sqrt{a} \mathbf{I}_2$, $\mathbf{V}_A = \mathbf{I}_2$, and $\mathbf{U}_A = \mathbf{A}/\sqrt{a}$.

-
- [1] GEO600 Project, www.geo600.uni-hannover.de; LIGO Collaboration, <http://www.ligo.org>; Virgo Collaboration, <http://www.virgo.infn.it>.
 - [2] M. Feo *et al.*, in *Proceedings of the 7th Marcel Grossmann Meeting*, edited by R. T. Jantzen *et al.* (World Scientific, Singapore, 1996), p. 1086.
 - [3] J.-M. Innocent and B. Torr sani, *Appl. Comput. Harmon. Anal.* **4**, 113 (1997).
 - [4] E. Chassande-Mottin and P. Flandrin, *Appl. Comput. Harmon. Anal.* **6**, 252 (1999).
 - [5] F.A. Jenet and T.A. Prince, *Phys. Rev. D* **62**, 122001 (2000).
 - [6] W. Anderson and R. Balasubramanian, *Phys. Rev. D* **60**, 102001 (1999).
 - [7] E. Cand s, P. Charlton, and H. Helgason, *Appl. Comput. Harmon. Anal.* **24**, 14 (2008).
 - [8] E. Chassande-Mottin and A. Pai, *Phys. Rev. D* **73**, 042003 (2006).
 - [9] N. Arnaud *et al.*, *Phys. Rev. D* **68**, 102001 (2003).
 - [10] H. Mukhopadhyay, N. Sago, H. Tagoshi, H. Takahashi, and N. Kanda, *Phys. Rev. D* **74**, 083005 (2006).
 - [11] A. Pai, S. Dhurandhar, and S. Bose, *Phys. Rev. D* **64**, 042004 (2001).
 - [12] L. S. Finn, *Phys. Rev. D* **63**, 102001 (2001).
 - [13] S. Klimenko, S. Mohanty, M. Rakhmanov, and G. Mitselmakher, *Phys. Rev. D* **72**, 122002 (2005).
 - [14] S.D. Mohanty, M. Rakhmanov, S. Klimenko, and G. Mitselmakher, *Classical Quantum Gravity* **23**, 4799 (2006).
 - [15] M. Rakhmanov, *Classical Quantum Gravity* **23**, S673 (2006).
 - [16] Y. G rsel and M. Tinto, *Phys. Rev. D* **40**, 3884 (1989).
 - [17] L. Wen and B.F. Schutz, *Classical Quantum Gravity* **22**, S1321 (2005).
 - [18] S. Chatterji *et al.*, *Phys. Rev. D* **74**, 082005 (2006).
 - [19] E.E. Flanagan and S.A. Hughes, *New J. Phys.* **7**, 204 (2005).
 - [20] K.S. Thorne, *Rev. Mod. Phys.* **52**, 299 (1980).
 - [21] E. Berti *et al.*, *Phys. Rev. D* **76**, 064034 (2007).
 - [22] E.E. Flanagan and S.A. Hughes, *Phys. Rev. D* **57**, 4535 (1998).
 - [23] H. Goldstein, *Classical Mechanics* (Addison-Wesley, Redwood City, CA, 1980), 2nd ed.
 - [24] G.H. Golub and C.F. VanLoan, *Matrix Computations* (Johns Hopkins University Press, Baltimore, 1996), 3rd ed.
 - [25] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis* (Cambridge University Press, Cambridge, England, 1991).
 - [26] S. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory* (Prentice-Hall, Englewood Cliffs, NJ, 1998).
 - [27] A. Neumaier, *SIAM Rev.* **40**, 636 (1998).
 - [28] J. Sylvestre, *Phys. Rev. D* **68**, 102005 (2003).
 - [29] H.L. Van Trees, *Optimum Array Processing* (Wiley, New York, 2002).
 - [30] K.R. Nayak, S.V. Dhurandhar, A. Pai, and J.Y. Vinet, *Phys. Rev. D* **68**, 122001 (2003).
 - [31] Y. G rsel and M. Tinto, in *Proceedings of General Relativity and Gravitational Physics, Capri, 1990* (unpublished), p. 620.
 - [32] E. Chassande-Mottin and A. Pai, *IEEE Signal Process. Lett.* **12**, 508 (2005).
 - [33] O. Rabaste, E. Chassande-Mottin, and A. Pai, in *Actes du 21 me Colloque GRETSI* (Troyes, France, 2007), p. 741.
 - [34] S.V. Dhurandhar and M. Tinto, *Mon. Not. R. Astron. Soc.* **234**, 663 (1988).
 - [35] I.M. Gel'fand, R.A. Minlos, and Z.Y. Shapiro, *Representations of the Rotation and Lorentz Groups and their Applications* (Pergamon Press, New York, 1963).