Gravitational Wave Polarizations:
a General Relativity Test

Samuel Patrone

INFN, Sezione di Roma, I-00185 Roma, Italy

Alan Weinstein (Mentor)

LIGO Laboratory, California Institute of Technology, Pasadena, California 91125, USA
(Dated: January 2, 2020)

This research aims at assessing with software simulations the possibility of inferring the polarization content of a transient gravitational waves (GWs) within a Bayesian framework, given the extended network of five ground interferometers available in the near future. The response of a network of GW interferometric detectors will be studied in order to discriminate between different polarizations. The methodology developed - on simulated data - provides a powerful check for one of the fundamental predictions of General Relativity, and is thus of major significance in both the fundamental physics of gravity and the astrophysics of gravitational-wave sources. In the following, it is shown that through Bayesian inference it is indeed possible (in simulation) to extract the polarization content of transient gravitational wave signals accurately, and with precision limited by the signal-to-noise ratio (SNR) in the response. The precision mostly improves with the network mean SNR as expected, with some exceptions, due either to a SNR distributed very unevenly between the five detectors or statistical fluctuations in the Bayesian sampler.

I. CONTEXT

Metric theories of gravity alternative to General Relativity (GR) imply the presence of specific GW polarizations other than the tensor ones predicted by GR. A fundamental goal of GW physics is therefore to find viable ways to detect and measure GW polarizations. We begin with a brief summary about detecting and studying GW polarizations with laser interferometric GW detectors. The notation used is taken from [1].

A. Linearized Metric Theory of Gravity

In the weak-field regime (or in the far-field limit), an appropriate coordinate system can be found in which the full metric $g_{\mu\nu}$ can be expressed as a perturbation $h_{\mu\nu}$ of order ϵ of the Minkowskian metric $\eta_{\mu\nu}$:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \,. \tag{1}$$

We can introduce a restricted class of coordinate transformations, called *Gauge Transformations*, such that the transformed metric is still separable into a flat *background* and a perturbation. These transformations are induced by a vector field $\zeta^\mu(x^\nu)$, with $\|\zeta^\mu_{,\nu}\| \sim O(\epsilon)$ such that the perturbation in the new coordinates becomes:

$$h'_{\mu\nu} = h_{\mu\nu} - \zeta_{\{\mu,\nu\}} + O(\epsilon^2) \,. \tag{2}$$

In terms of the perturbed metric, the *Riemann Tensor* is written as:

$$R_{\mu\nu\alpha\beta} = \frac{1}{2}(h_{\mu\beta,\nu\alpha} + h_{\nu\alpha,\mu\beta} - h_{\mu\nu,\alpha\beta} - h_{\alpha\beta,\mu\nu}) \,. \tag{3}$$

It is noteworthy to stress that the Riemann Tensor is an invariant under gauge transformations, since it encodes only the information about gravity, without keeping track of the chosen coordinate system.

Fixing a gauge is a way to use the freedom we have in the choice of the vector field ζ^μ to simplify our problem and reduce the ten degrees of freedom of the symmetric rank-2 tensor $h_{\mu\nu}$ down to six. In GR, a common gauge choice in vacuum is the Lorentz gauge (also known as *harmonic gauge*) $\bar{h}^\mu_{\nu,\mu} = 0$, where $\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}h\eta_{\mu\nu}$. This gauge is adopted because it simplifies the Einstein tensor. The residual gauge freedom given by field equations (direct consequence of the *Bianchi identity* applied to the Einstein tensor) allows one to further impose $h_{0i} = 0$ and $h = 0$, leaving only two radiative degrees of freedom, the two *tensorial* polarizations. This last choice is often called the *trasverse-traceless gauge*, or TT-gauge. In a non-GR theory, since the field equations are different, we have to take into account all six degrees of freedom.

B. GW Polarizations

We want to describe the Fabry-Perot Michelson interferometer hit by a (plane) GW as two couples of test masses (at the edge of the two Fabry-Perot cavities, situated in the two arms of the detector) subject to a metric perturbation: the response of the detector depends on the difference in travel time along the two arms. The behavior of the separation vector ξ^α between the two freely-falling test masses with 4-velocity u^β is given by the equation of geodesic deviation:

$$\frac{D^2\xi^\mu}{dt^2} = -R^\mu_{\nu\alpha\beta}u^\nu\xi^\alpha u^\beta \,. \tag{4}$$
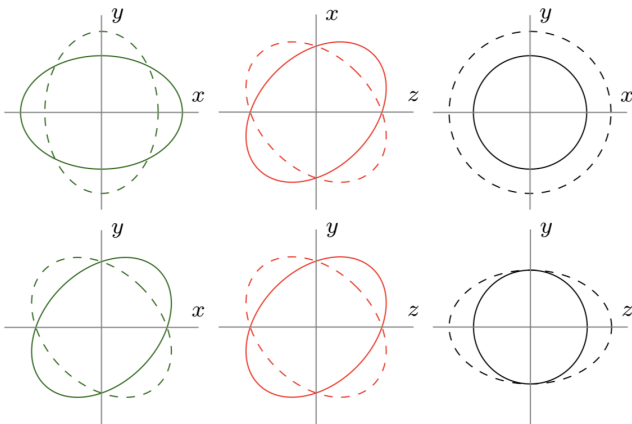
FIG. 1. *Effect of different GW polarizations on a ring of free-falling test particles.* Plus (+) and cross (×) tensor modes (green); vector-x (x) and vector-y (y) modes (red); breathing (b) and longitudinal (l) scalar modes (black). In all of these diagrams the wave propagates in the z direction. This decomposition into polarizations was first proposed for generic metric theories in [2]. (Reproduced from [3].)

For slowly moving particles, we can rewrite the previous equation as:

$$\frac{d^2\xi_j}{dt^2} = -c^2 R_{0j0k}\xi^k = -\frac{G}{2c^4 D}\frac{\partial}{\partial\tau^2} S_{jk}(\tau, \mathbf{N})\xi_k. \quad (5)$$

where D is the distance from the source of GW, $\mathbf{N} = (\cos\phi\sin\theta, \sin\phi\sin\theta, \cos\theta)$ is a unit 3-vector which points toward the direction of the source in the sky (direction of propagation of the wave) while the S_{ij} contains the proper time dependent amplitudes of the perturbation. Integrating over time, we obtain the equation of motion for the displacement of the detector arms at first order:

$$\xi^j(t) = \xi^j(0) + \frac{G}{2c^4 D}S^{jk}(\tau, \mathbf{N})\xi_k(0). \quad (6)$$

Assuming that the wave travels along the z-direction, we have:

$$S_{jk} = \begin{pmatrix} A_b + A_+ & A_\times & A_{Vx} \\ A_\times & A_b - A_+ & A_{Vy} \\ A_{Vx} & A_{Vy} & A_l \end{pmatrix}. \quad (7)$$

Six degrees of polarization can be identified through their effect on a ring of free-falling test particles (see fig. 1). Two scalar (A_b and A_l), called respectively *breathing* and *longitudinal* modes, one of which (A_b) is transverse with respect to the direction of propagation of the wave. Two vector (A_x and A_y) modes, partly longitudinal and partly transverse. Finally, two tensorial ($A_\times$ and A_+) transverse modes. GR only allows for $A_\times$ and A_+.

From a field-theoretic point of view, polarizations are strictly correlated with the helicity (projection of the spin along the motion) of the *graviton*: a massless graviton has only ±2 helicity, which corresponds to the two tensorial polarizations of GR.

II. STUDY OF THE SENSITIVITY OF A NETWORK OF INTERFEROMETRIC DETECTORS

A. Antenna patterns of a two-arm and three-arm interferometer

While the amplitudes and the phases of GWs depend crucially on the source dynamic, the response of a quadrupolar antenna to them is determined by the geometry of the system source-detector (up to an overall normalization). Let $\mathbf{e_1}$ and $\mathbf{e_2}$ be the unit vectors aligned with the two arms of the interferometer. The *differential arm* (DARM) response function is given by:

$$S(t) = \frac{1}{2}(e_1^j e_1^k - e_2^j e_2^k)S_{jk}(\tau, \mathbf{N}). \quad (8)$$

We can now consider an orthonormal basis $(\mathbf{N}, \mathbf{e_X}, \mathbf{e_Y})$ to fully describe the polarization of the wave, where $(\mathbf{e_X}, \mathbf{e_Y})$ are rotated by an angle ψ (called *angle of polarization*) around $\mathbf{N}$ with respect to the basis $(\mathbf{e_1'}, \mathbf{e_2'})$. These primed vectors are obtained by two subsequent rotations of $(\mathbf{e_1}, \mathbf{e_2})$ that align the basis of the interferometer along $\mathbf{N}$. We can write the total response as a function of three angles (θ, ϕ, ψ):

$$S(t) = F_P(\theta, \phi, \psi)A_P \quad (9)$$

where $P = \{b, l, x, y, +, \times\}$, the F's are called *antenna pattern functions* and the sum runs over all polarizations P.

Preliminarily, a complete study of the antenna patterns for a two-arm and a three-arm interferometer was performed. The results are reported in fig.2 and fig.3. For a detailed description of the three-arm interferometer simulated, see [4]. As it is shown, the detector response to the scalar polarizations (apart from a sign) is completely degenerate for both the configurations, therefore the two cannot be distinguished from each other and from now on we will consider only a single scalar mode S, with corresponding antenna response F_S. Furthermore, the sensitivity of a three-arm detector is - on average - greater than the two-arm configuration, presenting also a peculiar azimuthal symmetry in the detector frame.

B. Network sensitivity and overlap factor

In the measurement of GW polarizations, eight different unknowns play a role: the six polarization modes and the two angles that identify the position of the source in the sky. However, as mentioned earlier, the DARM response functions of a laser interferometer to the two scalar modes are completely degenerate, and we are therefore left with five independent polarization modes. The number of available quadrupolar antennas is then crucial to measure the polarization content. Furthermore, the orientation of the arms of the instruments
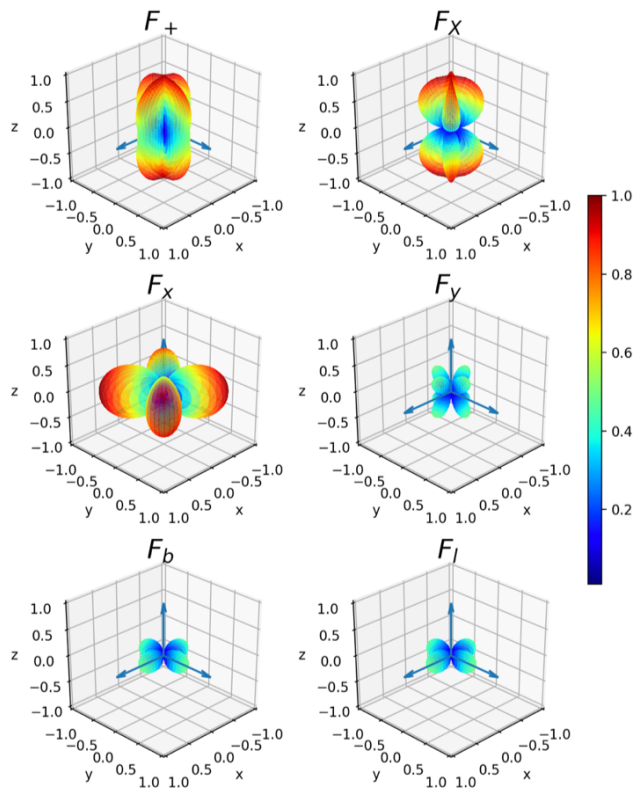
FIG. 2. *Angular response of a two-arm interferometer to each GW polarization.* The radial distance represents the response of a single two-arm quadrupolar antenna to a unit-amplitude gravitational signal of a tensor (top), vector (middle), or scalar (bottom) polarization, i.e. $|F_P|$ for each polarization P for $\psi = 0$. The polar and azimuthal coordinates correspond to the source location with respect to the detector, which is to be imagined as placed with its vertex at the center of each plot and arms along the x and y-axes. The response is plotted to scale, such that the blue lines representing the detector arms have unit length in all plots.

plays a fundamental role. If the arms of a pair of interferometers are aligned, while the the probability of coincident (between detectors) detection is maximized, there is the downside that the antenna pattern function will be the same, preventing the possibility of distinguishing between different polarizations. Over the next ten years, two new interferometers will be available: KAGRA (in $\sim$ 2020) and LIGO-India. Additionally, the prospect of building a new generation of detectors (so called 3G detectors) is under discussion and investigation (for further references, see [5] and [6]). In this study, the response of the five-interferometer network is studied, which includes the following detector: LIGO Hanford (USA), LIGO Livingston (USA), Virgo (Italy), KAGRA (Japan), and LIGO-India. The position of each detector on the Earth's surface is shown in fig.4.

If we are interested to the sensitivity of a network of N detectors, setting $\psi = 0$ since we are not dealing with any specific source, it is useful to define the effective response
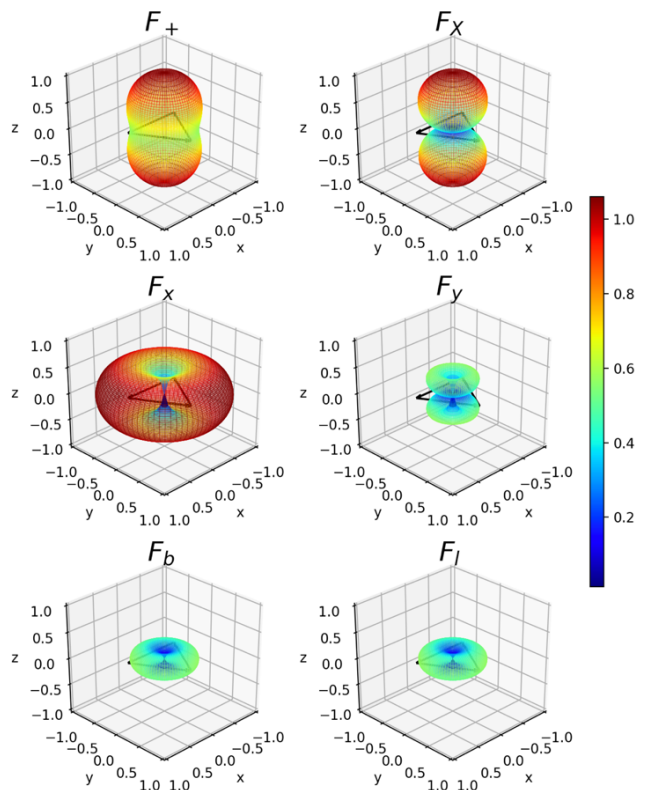

FIG. 3. *Angular response of a three-arm interferometer to each GW polarization.* The radial distance represents the response of a system with a triangular topology, where the arms of the equilateral triangle are each used twice to form three Michelson interferometers. It can be analytically shown that the response to GW different polarizations is equivalent to that of two L-shaped detectors of length 3L/4. In the figure, the response to a unit-amplitude gravitational signal of a tensor (top), vector (middle), or scalar (bottom) polarization is shown, i.e. $|F_P|$ for each polarization P for $\psi = 0$. The polar and azimuthal coordinates correspond to the source location with respect to the detector, which is to be imagined as placed with its center in the origin of each frame. The response is plotted to scale, such that the black lines representing the detector arms have unit length in all plots.
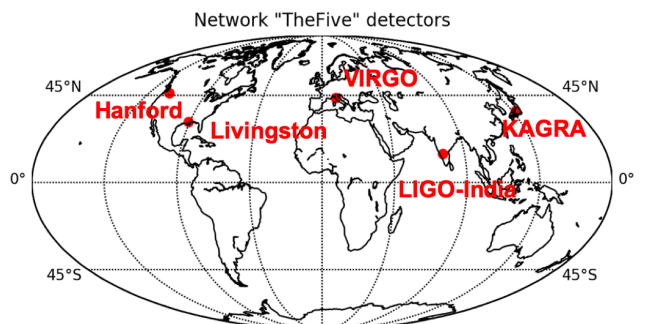

FIG. 4. *Five-interferometer network available in the near future, used for this research.*
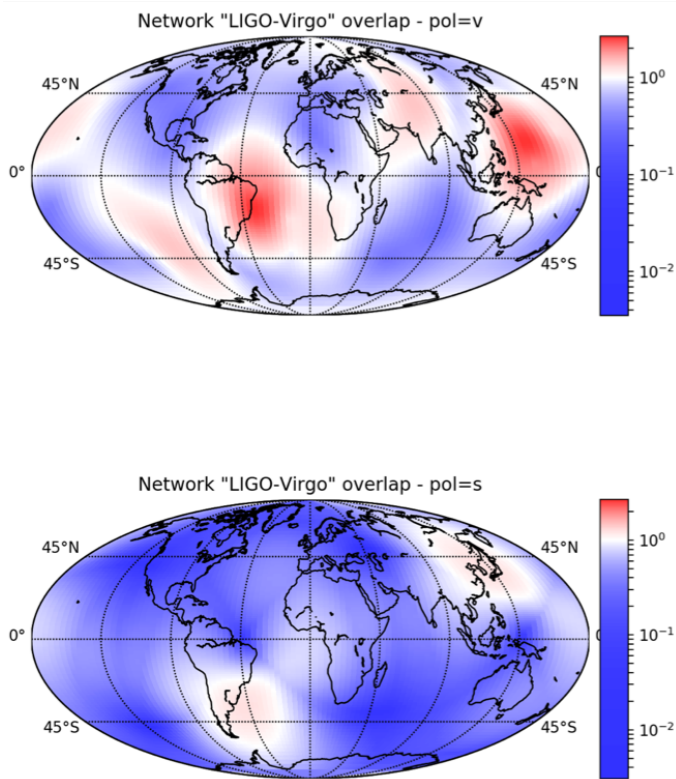
FIG. 5. *Overlaps of LIGO-Virgo network effective antenna patterns.* The normalized inner-products of Eq. (11) for the three-instrument network. The top plot compares vector to tensor $\mathcal{F}_{v/t}$, and the bottom one compares scalar to tensor $\mathcal{F}_{s/t}$. Red (blue) marks regions for which the effective non-tensor response is greater (less) than tensor. A map of Earth is overlaid for reference.

FIG. 6. *Overlaps of the five-detector network effective antenna patterns.* The normalized inner-products of Eq. (11) for the five-instrument network. The top plot compares vector to tensor $\mathcal{F}_{v/t}$, and the bottom one compares scalar to tensor $\mathcal{F}_{s/t}$. Red (blue) marks regions for which the effective non-tensor response is greater (less) than tensor. A map of Earth is overlaid for reference.

vector as:

$$\vec{F}_H(\theta, \phi) := (|F_H^1(\theta, \phi)|, ..., |F_H^N(\theta, \phi)|) \qquad (10)$$

where the F_H^i's are the sum in quadrature of the two *antenna patterns* of the i-th detector for each polarization $H = \{s, v, t\}$ (scalar, vector, tensor). We can evaluate the effective sensitivity of the network to non-tensorial polarizations with respect to tensorial ones by computing the *overlap* factor:

$$\mathcal{F}_{H/t} = \frac{\vec{F}_H(\theta, \phi) \cdot \vec{F}_t(\theta, \phi)}{\vec{F}_t(\theta, \phi) \cdot \vec{F}_t(\theta, \phi)}. \qquad (11)$$

It is essential to quantify in advance, through simulations, how GW detector configuration choices affect our ability to measure the GW polarization content, as this measurement can place strong, fundamental constraints on theories of gravity. From the definition of *overlap* given in Eq. (11), we can draw skymaps of relative sensitivity. This has been done in [3] for the three-detector LIGO-Virgo network. We extended the study up to the five-interferometer network.

To allow the comparison with the actual three-interferometer network, skymaps for both networks are
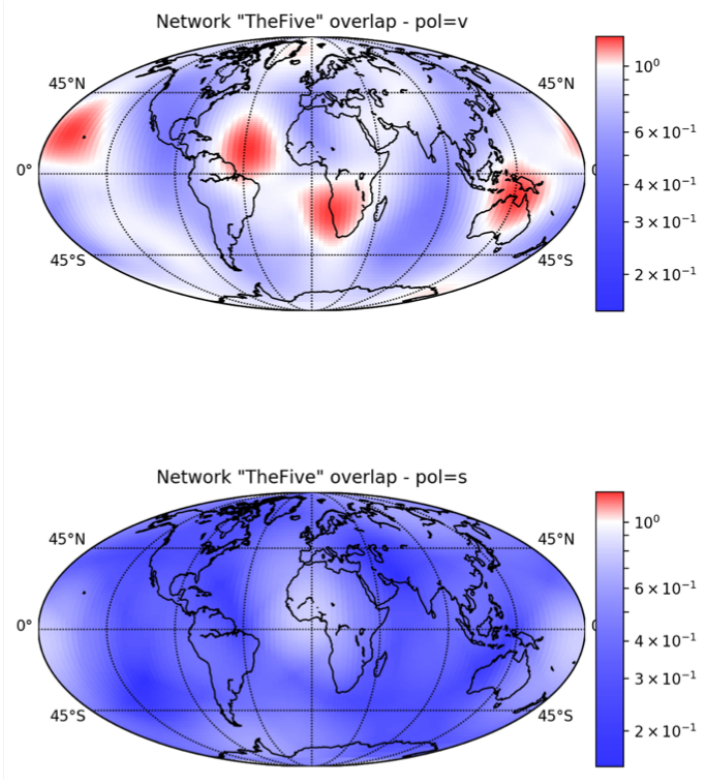
shown, in fig.5 (LIGO-Virgo network) and fig.6 (five-detector network). Colored regions roughly correspond to areas in the sky for which the tensor and the non-tensor responses are highly distinguishable. Notice that the patterns are anchored to angular locations with respect to Earth (not the fixed stars): in other terms, the shown sensitivity depends only on the the specific location and orientation of the detectors in the considered network. To find the correspondent sky location in the celestial sphere, sidereal time of arrival of the signal must be taken into account.

For a more detailed quantitative study of the overlap factor, the probability distribution function (for all sky locations) of each overlap is plotted in fig.7 and fig.8. The extended network (that from now on will be referred as *the network*) will be the least sensitive to scalar modes (since the interferometers are individually less sensitive to these); but with five detectors, the sensitivity to scalar polarization is improved by one order of magnitude with respect to the three-detector configuration. From this first quantitative analysis, which is exclusively based on the reciprocal position and orientation of detectors, it is already possible to infer that with the addition of two
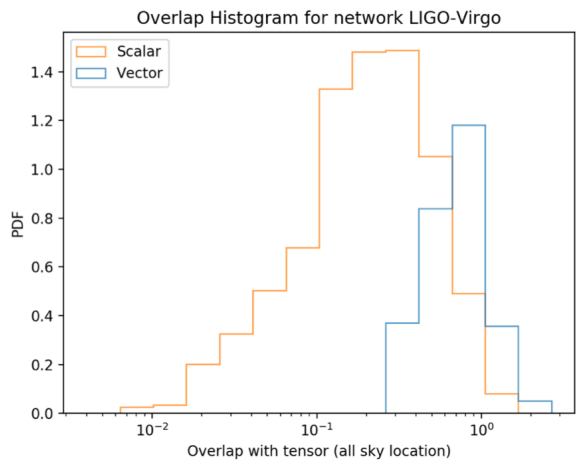
FIG. 7. *Probability distribution function of the overlaps for LIGO-Virgo network.*
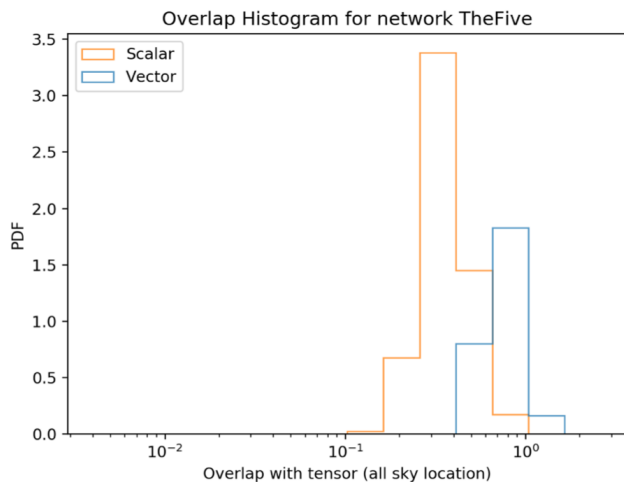

FIG. 8. *Probability distribution function of the overlaps for the five-detector network.*

interferometers, the response to non-GR polarization is mostly improved.

III. SIMULATED RESPONSE OF THE NETWORK

A. Methods to extract polarization content

As we have seen previously (Sec. II A), antenna patterns are a direct manifestation of local geometry only (polarizations and detector geometry), independent of source or the details of the underlying theory. We may thus exploit the difference in the response of the network to the different polarizations to infer the polarization content of the wave.

There are two ways to extract polarization from antenna patterns if the GW signal is transient.

In the first scenario in which the GW signal has an optical counterpart that allows for an accurate determination of the source position [7], it is particularly convenient to look for non-GR signal content by constructing *GR null stream(s)*, which are *directions* in the multidimensional space of the network response (whose dimensions are given by the total number of detectors) in which there should be no signal if the wave is tensor-only polarized. Indeed, for N detectors, the signal manifold is N-dimensional with N basis vectors, five of which can be chosen along the independent antenna patterns $F_+^i, F_\times^i, F_x^i, F_y^i, F_s^i$ where the Latin index runs along the N detectors. The remaining $N-5$ vectors will give us null streams, independently from the polarizations of the wave. The j-th detector datastream can be written in tensor notation as:

$$S^j = F_P^j h^P + n^j , \qquad (12)$$

where n^j is the noise content in the j-th detector. In the case of three detectors, we can define a *GR null stream*, i.e., a stream without tensor modes, in the following way (see [8]):

$$S_{GR-null} = \frac{e_{ijk} F_+^j F_\times^k}{|\delta_{ij} F_+^i F_\times^j|} S^i . \qquad (13)$$

Depending on the number of interferometers, we can have more than one null stream, and with more than five, one can construct a complete set of null streams that covers all metric theories of gravity. This method is model independent, but it has the disadvantage of requiring an electromagnetic counterpart.

A second method which doesn't require necessarily an electromagnetic counterpart is the following: using a sine-Gaussian analysis to reconstruct the waveform, one may infer from time delays the source location and then the best fitting combination of antenna patterns for the peak in amplitude. This analysis is independent from the phase evolution and it only needs a well-defined peak (as shown in [3], Sec. III A). With three interferometers, it is already possible to infer the direction $\mathbf{N}$ of the source in the sky just measuring time delays, which is given by the formula:

$$\delta t_I = \mathbf{N} \cdot \mathbf{x}_I / c . \qquad (14)$$

where δt_I is the time delay with respect to the geocenter and $\mathbf{x}_I$ joins the geocenter to the detector. With four interferometers, constraints on the propagation velocity of GWs can be placed, providing information about the mass of the graviton and, indirectly, on GW polarizations. The key idea is that, in such test, *no polarization information is extracted from the phase evolution*: in other terms, the template is only used to infer the source location from the time lag between detectors, and the best-fitting combination of antenna patterns from the amplitudes and phases at peak energy. We chose to adopt this methodology for our research.

B. Sine-gaussian toy model for a generic transient Gravitational Wave

A specific source of gravitational waves (with non-GR polarizations) is generically identified by the following astrophysical parameters:

- Sky coordinate location: (δ, α);

- Polarization angle: ψ;

- Geocentric sidereal time of arrival: t_0;

- Luminosity distance of the source: d_L.

In the context of the second methodology described in the previous subsection (Sec. III A), since we don't know the actual waveform of a non-GR polarized wave, it is reasonable to build a simplified toy model of the response at each two-arm detector I using a single sine-gaussian wavepacket with given frequency Ω and relaxation time τ:

$$h_I(t) = A_I \cos(\Omega t + \phi_I) \exp\left[-\frac{(t - t_0 - \delta t_I)^2}{2\tau^2}\right], \quad (15)$$

$$A_I = \frac{|A|}{d_L}|\tilde{A}_I|, \quad (16)$$

$$\tilde{A}_I = \sum_{p \in \{+, \times, x, y, s\}} \epsilon_p F_p^I. \quad (17)$$

where:

- $\epsilon_p = \frac{a_p}{|A|} e^{i\phi_p}$ are complex coefficients which depend exclusively on the unknown underlying properties of the GW generation mechanism (e.g. the tensor ones, within a triaxial emission model, are functions of the inclination angle ι) and they are also coordinate frame dependent quantities. Since we pick $|A|$ to be the total amplitude, we expect $|\epsilon_p| \in [0, 1]$;

- $|A| = \sqrt{\sum_p a_p^2}$ is the overall amplitude of the wave;

- $F_p^I = F_p^I(\alpha, \delta, \psi = 0, t_0)$ are the antenna patterns for the detector I. Through setting $\psi = 0$, we are choosing a specific polarization frame: we can arbitrarily fix the polarization angle thanks to the degeneracy between ψ and the two couples of parameters $(\epsilon_+, \epsilon_\times)$ and (ϵ_x, ϵ_y), consequence of the rotational property of antenna patterns (see Appendix A in [9]);

- $\phi_I = \arctan\frac{\mathcal{I}[\tilde{A}_I]}{\mathcal{R}[\tilde{A}_I]} - \Omega(t_0 + \delta t_I)$ is the measured phase at the I detector.
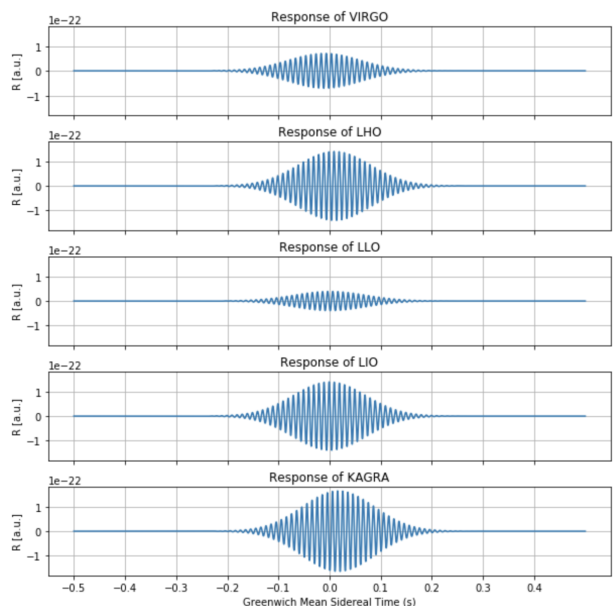

FIG. 9. *Example of a simulated GW signal as seen by the network.* Concerning the astrophysical parameters, the signal shown was generated with random sky location (α, δ), random polarization angle ψ, sidereal time of arrival at the center of the earth $t_0 = 0\,s$, unitary luminosity distance $d_L = 1$. Both lambdas were set to 0.1 with fixed overall amplitude $|A| = 1e - 22$, randomly choosing the remaining amplitude parameters (the two ϵ's and the five phases ϕ_p). The total duration of the segment is $1\,s$ and the sine-gaussian was chosen with frequency $\Omega = (2\pi)100\,Hz$ and damping time $\tau = 0.1\,s$. Time shifts between signals of several tens of ms are due to the signal propagation (assumed at the speed of light) through the Earth, while amplitude changes are a direct consequence of different detectors' antenna response functions (see II A). It's worthy to stress that the polarization analysis is based on these last amplitude (and phase) differences.

We can also define the *normalized effective strain amplitudes* in the following way:

$$h_t = \sqrt{|\epsilon_+|^2 + |\epsilon_\times|^2}, \quad (18)$$

$$h_v = \sqrt{|\epsilon_x|^2 + |\epsilon_y|^2}, \quad (19)$$

$$h_s = |\epsilon_s|. \quad (20)$$

We can introduce two hyper-parameters (coordinate independent) to quantify how much of the wave is tensor, vector or scalar polarized, λ_v and λ_s, such that:

$$h_t^2 = 1 - \lambda_v - \lambda_s, \quad (21)$$

$$h_v^2 = \lambda_v, \quad (22)$$

$$h_s^2 = \lambda_s \,. \qquad (23)$$

The main goal of this research is to establish how well we can infer the two lambda hyper-parameters from signals with different SNR. An example of the signal as seen in each detector of the network can be found in fig.9.

C. Noise, matched filtering and SNR

To give a realistic simulation of the network response, a gaussian colored noise (weighted in frequency space) was added to the signal timeseries.

We can imagine our response $d(t)$ at each detector I as a superposition of a pure signal $h(t)$ and a gaussian colored noise $n(t)$, with a different variance $\sigma(f)$ according to a specific power spectral density $S_n(f)$ of the noise for each detector. In formulae:

$$d_I(t) = h_I(t) + n_I(t) \,, \qquad (24)$$

$$S_n^I(f) = \lim_{T \to \infty} \frac{2}{T} \left| \int_{-T/2}^{T/2} n_I(t) e^{2\pi i f t} dt \right|^2 \,. \qquad (25)$$

In this simulation, we used the designed analytical power spectral density $S_n(f)$ for each detector, in the most optimistic prevision. If we define the functional W as the following:

$$W = \int_{-\infty}^{\infty} d(t) K(t) dt \,, \qquad (26)$$

it can be shown that W will be maximally sensitive to the signal if we use as a kernel filter the inverse fourier trasform of the following:

$$\tilde{K}(f) = \frac{4\tilde{h}(f)}{S_n(f)} \,, \qquad (27)$$

where $\tilde{h}(f)$ is the Fourier transform of the signal template.

This last object is called *Wiener's optimal filter*. In virtue of Parseval's theorem, we can compute norms integral both in time and in frequency space, obtaining the same result (since Fourier transform is a unitary linear operator). We can now define the Signal-to-Noise Ratio (SNR) to be:

$$\rho(t) = \frac{\langle h|d \rangle}{\sqrt{\langle h|h \rangle}} \,, \qquad (28)$$

where we used the notation $\langle h|s \rangle$ for the following scalar product ($\mathcal{R}$ stands for the real part)

$$\langle h|d \rangle = 4\mathcal{R} \int_0^{\infty} \frac{\tilde{h}^*(f)\tilde{d}(f)e^{2\pi i f t}}{S_n(f)} df \,. \qquad (29)$$

As it can be seen from the previous relations, the SNR is a function of time and it can be used for a measure of how well we can distinguish the signal (and then extrapolate its waveform) from a noisy background. By Eq. (29), it seems necessary to know in advance the *template $h(t)$* for the signal to recover the correct SNR (and then to see the signal itself). However, it is worthy to point out that this analysis can be carried out also in a model-independent way, using for example a basis of sine-gaussian waveforms to reconstruct the signal (see [10]).

An example of a noisy response is shown in fig.10 and the SNR time series $\rho(t)$ is shown in fig.11.

In the following, we will call the maximum of the function $\rho(t)$ for each signal the SNR of the event. Notice that for a pure noise signal, $\rho(t)$ oscillates in the interval $[0, 3]$. Usually, an event is considered a signal when it has a SNR bigger than 8. Noise generation and matched filtering algorithm were both implemented using pyCBC libraries [11].
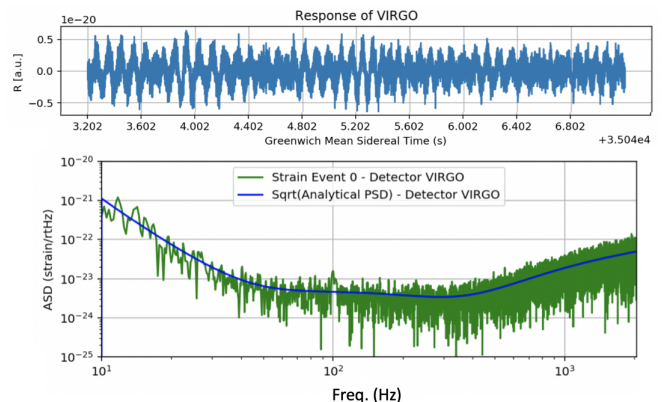

FIG. 10. *Example of a simulated GW noisy response of VIRGO detector and its strain in frequency space.* The signal is not distinguishable by eyes in time space, but its strain in frequency space (the square root of the PSD of the full response) reported in the second plot shows a little bump at $100\,Hz$, the frequency chosen for the sine-gaussian. For comparison, it is over plotted the square root of the analytical PSD used to generate the noise. Notice that, since the total length of the signal in time is $4\,s$, the Discrete Fourier Transform (DFT) has a maximum resolution in frequency of $T^{-1} = 0.25\,Hz$.
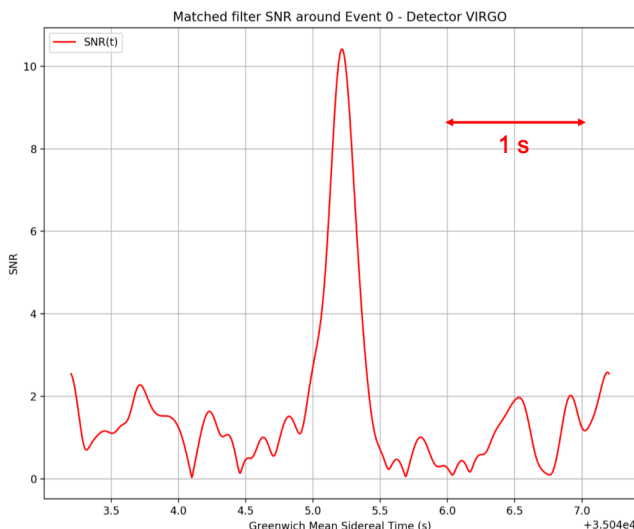
FIG. 11. *Time evolution of the Signal-to-Noise Ratio of the response in fig.10.* As expected, the SNR shows a peak for the time of arrival of the wave. Since its maximum is barely above 10, we can classify this event as a low SNR one.

IV. BAYESIAN FRAMEWORK

A. Context

We can divide the task of using Bayes Theorem to analyze a set of data $\mathbf{D}$ into two problems: parameters estimation and model selection.

Parameters estimation. Given a specific model $\mathcal{H}$ of our data which depend on a set of parameters $\boldsymbol{\theta}$, we can recover the joint *posterior* probability distribution function on the multidimensional space of parameters $P(\boldsymbol{\theta}|\mathcal{H}, \mathbf{D})$ through the following relation:

$$P(\boldsymbol{\theta}|\mathcal{H}, \mathbf{D}) = \frac{P(\boldsymbol{\theta}|\mathcal{H})P(\mathbf{D}|\mathcal{H}, \boldsymbol{\theta})}{P(\mathbf{D}|\mathcal{H})} \qquad (30)$$

This is just a trivial application of Bayes Theorem, where:

- $P(\boldsymbol{\theta}|\mathcal{H})$ are the *prior* of our model, joint probability distribution function that weighs the parameter space differently according to our previous knowledge on the phenomenon;

- $P(\mathbf{D}|\mathcal{H}, \boldsymbol{\theta})$ is the *likelihood*, the core of the computation, a measure on how well our model with chosen parameters can describe the data observed;

- $P(\mathbf{D}|\mathcal{H}) = \int_{\boldsymbol{\Theta}} P(\mathbf{D}|\mathcal{H}, \boldsymbol{\theta})P(\boldsymbol{\theta}|\mathcal{H})d\boldsymbol{\theta}$ is the *evidence*, the likelihood integrated over all parameters space (weighted by the prior), which in this context works like a normalization factor, but which is crucial in model selection, as we will see in the following.

Model selection. We can also use Bayes theorem to give a statistic significance to a model compared with another,

deciding which of several models is more probable in light of the observed data $\mathbf{D}$, and by how much. This can be done very easily by considering the ratio of posterior probabilities obtained using two different model $\mathcal{H}_i$ and $\mathcal{H}_j$, defining the *odds* for $\mathcal{H}_i$ versus $\mathcal{H}_j$ as:

$$\mathcal{O}_j^i = \frac{P(\mathcal{H}_i)}{P(\mathcal{H}_j)} \frac{P(\mathbf{D}|\mathcal{H}_i)}{P(\mathbf{D}|\mathcal{H}_j)} = \frac{P(\mathcal{H}_i)}{P(\mathcal{H}_j)} \mathcal{B}_j^i \qquad (31)$$

In the last equation, we introduced the definition of *Bayes factor* $\mathcal{B}_j^i$ which is the ratio of the two evidence integrals.

B. Analysis

Summarizing the work carried out in [9], given a vector of responses $\mathbf{D}$ of our network, we want to test seven possible Bayesian hypotheses regarding its polarization content: it is purely tensor ($\mathcal{H}_t$), purely vector ($\mathcal{H}_v$), purely scalar ($\mathcal{H}_s$), scalar-tensor ($\mathcal{H}_{st}$), vector-tensor ($\mathcal{H}_{vt}$), scalar-vector ($\mathcal{H}_{sv}$), scalar-vector-tensor ($\mathcal{H}_{svt}$). Using Bayes theorem, we can expand the probability $P(\mathcal{H}_S|\mathbf{D})$ that, given the data, a signal hypotheses can be accepted:

$$P(\mathcal{H}_S|\mathbf{D}) = \sum_m P(\mathcal{H}_m)P(\mathbf{D}|\mathcal{H}_m)/P(\mathbf{D}), \qquad (32)$$

where $m \in \{t, v, s, st, vt, sv, svt\}$, $P(\mathcal{H}_m)$ is a prior on the model, $P(\mathbf{D}|\mathcal{H}_m)$ is the marginalized likelihood and $P(\mathbf{D})$ is a overall normalization constant. We can define then the *odds* to detect a non pure-tensor signal as:

$$\mathcal{O}_t^{nt} = \sum_{m \neq t} \frac{P(\mathcal{H}_m)}{P(\mathcal{H}_t)} \mathcal{B}_t^m \qquad (33)$$

There is a subtle difference between a generic tensor signal and a GR signal, which is relevant only if we know *a priori* the inclination of the source within the tri-axial GR model. An extended study of this difference can be found in Appendix A of [9].

Besides choosing between different models, we can use Bayesian statistics to obtain posterior probability density function on the parameters of a given template for the waveform, provided that we define correctly all the contributions that appear in Eq. (30).

Parameters space. Assuming the toy model as described in Sec. III B, fixing the sine-gaussian frequency Ω and damping time τ, we are left with a total of 12 independent parameters $\{\alpha, \delta, |A|, \phi_+, \phi_\times, \phi_x, \phi_y, \phi_s, \epsilon_\times, \epsilon_y, \lambda_v, \lambda_s\}$. Notice that d_L is degenerate with $|A|$, ψ is degenerate with some of the amplitudes (as shown in appendix A in [9]) and the time of arrival t_0 cannot be inferred since we simulate the response of the network for a symmetric interval in time with respect to t_0 itself ($2\,s$ before and $2\,s$ after). Since the heavy computational cost of each simulation,

we chose to reduce the dimensions of parameters space down to 6, assuming the overall amplitude and the five amplitude phases as known.

Priors. Priors in the sky locations are set to be uniform over the celestial sphere: the right ascension (α) prior is flat in the interval $[0, 2\pi)$ while the cosine of declination $(\cos \delta)$ prior is flat in the interval $[-1, 1)$. Priors on the amplitude parameters (the two $\lambda's$ and the two $\epsilon's$) are taken flat in the interval $[0, 1)$, provided we get the normalization to one of the five $\epsilon's$ square sum. Indeed, it can happen that these four amplitudes parameters are not compatible with each other: in this case, we reject the point in the parameter space, weighting it by a negative infinite likelihood, which is well handled by Python code as it can be implemented as a NumPy constant.

Likelihood. For each detector, we assume the noise (as described in Sec. III C) to be both *stationary* (the power spectral density $S_n(f)$ is constant in time) and *gaussian* in each frequency bin, characterized only by having zero mean and a known variance (estimated from $S_n(f)$ itself). Then the *likelihood* function for the model $\mathcal{H}$ of parameters $\boldsymbol{\theta}$, given the detector response as in Eq. (24), is simply the product of Gaussian distributions in each frequency bin i with adjusted mean value due to the presence of the signal:

$$
\begin{aligned}
P(d|\mathcal{H}, \boldsymbol{\theta}; S_n(f)) &= \exp \sum_i \left[-\frac{2 \left| \tilde{h}_i(\boldsymbol{\theta}) - \tilde{d}_i \right|^2}{T S_n(f_i)} \right. \\
&\left. -\frac{1}{2} \log\left(\pi T S_n(f_i)/2 \right) \right],
\end{aligned}
\tag{34}
$$

where

$$
\tilde{d}_j(f) = \frac{T}{N} \sum_k d_k e^{-2\pi i j k f} .
\tag{35}
$$

is the Discrete Fourier Transform (DFT) of the response (the index k runs over time and depends on the sampling frequency chosen, N is the total number of samples, T is the total length in time of the response) and $\tilde{h}$ is the DFT of the template, at given parameters $\boldsymbol{\theta}$.

To analyze a network of detectors coherently, we make the further assumption that the noise is uncorrelated in each. This allows us to write the coherent network likelihood for data obtained from each detector as the product of the likelihoods in each detector I:

$$
P(\mathbf{d}|\mathcal{H}, \boldsymbol{\theta}) = \prod_I P(d_I|\mathcal{H}, \boldsymbol{\theta}; S_n^I(f))
\tag{36}
$$

A graphical study on the mono-dimensional and bi-dimensional behaviour of the likelihood in parameter space is reported in Appendix A for a random signal.

Sampler. At last, we have to define an efficient way to explore parameters space, the so-called *Bayesian sampler*, an algorithm that progressively maximizes the likelihood, avoiding local maxima. The algorithm chosen was Skilling's *nested-sampling* [12]: the inference code used was built in the Dynesty package [13], a free Pure Python, MIT-licensed Dynamic Nested Sampling package for estimating Bayesian posteriors and evidences. The key idea of nested-sampling is to numerically compute the evidence, by reducing it into a one-dimensional integral, defining the variable change in a proper way. Posteriors on parameters can then be derived with a minimal computational cost. For further details and its application to GW signals with a detector network, see [14].

V. RESULTS

In this section, we present the results of our research. As a first example to validate our methodology, we tested our inference code on two signals of different SNR, to recover six parameters of the wave (sky location and four amplitudes parameters), as described in Sec. IV B. Trace plots and corner plots of the results are reported in fig.12 and in fig.13. Trace plots show the evolution of the sampled points in parameters space as a function of the variable X, which corresponds to the one-dimensional variable used to compute the evidence integral. This variable (or more properly, $-\ln(X)$) can be seen as a proxy to the number of iterations the algorithm needs to converge to the global maximum of the likelihood. Corner plots show the two-dimensional posteriors (marginalized over all other parameters) for each parameters pair and the mono-dimensional marginalized posterior distribution of each single parameter on the diagonal, as recovered by the algorithm. Red lines identify the injected value in each plot.

In both cases, the injected value is recovered correctly with an high accuracy (generally within one sigma), but the precision of the result crucially depends on the SNR. In the second simulation, in which the SNR is much lower, the posterior distributions are much wider and less peaked on the true value. This is the first important result of our study: the network (within this model) is able to disentangle all the polarization degrees of the wave, but the precision (i.e. the sensitivity of the network to each specific polarization) strictly depends on the strength of the signal and its SNR in each detector.

To further analyze the systematic of this behaviour, we simulated several different signals from sources with increasing value of the lambdas hyper-parameters, randomly setting the other ten parameters to have for each value of lambda, three different SNR (in each detector): greater than 100, between 25 and 100, lower than 30. We repeated this study for a mixture of two polarizations (setting one of the two lambdas to zero) and three polarizations (allowing both lambdas to be different from zero and set to the same value). The marginalized pos-
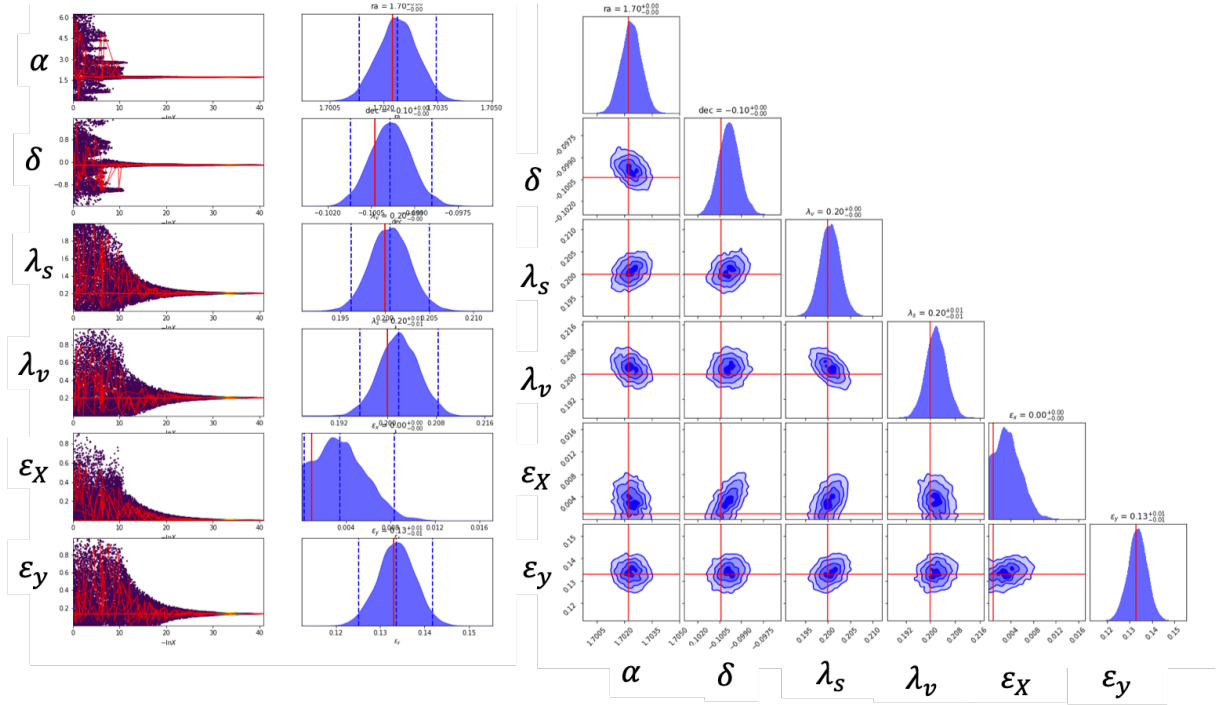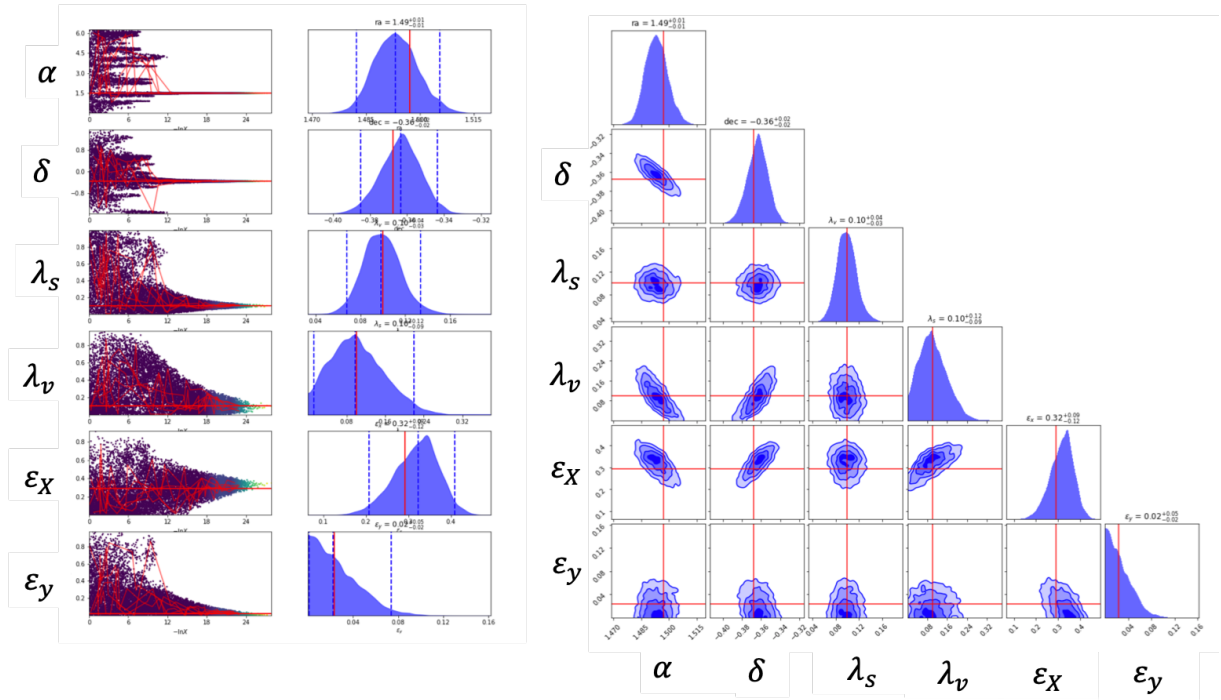
FIG. 12. *Trace plots and corner plots of a high SNR source.* The injected true value (red lines in both plots) is well recovered after some iterations of the sampling algorithm, with high accuracy and high precision. This specific source was randomly chosen with an $\epsilon_\times$ close to zero: the algorithm correctly provides upper limits for its value.
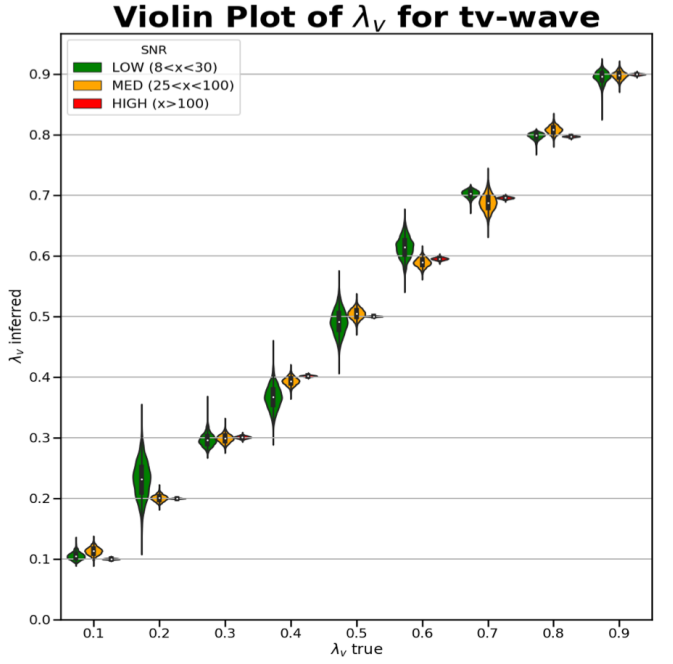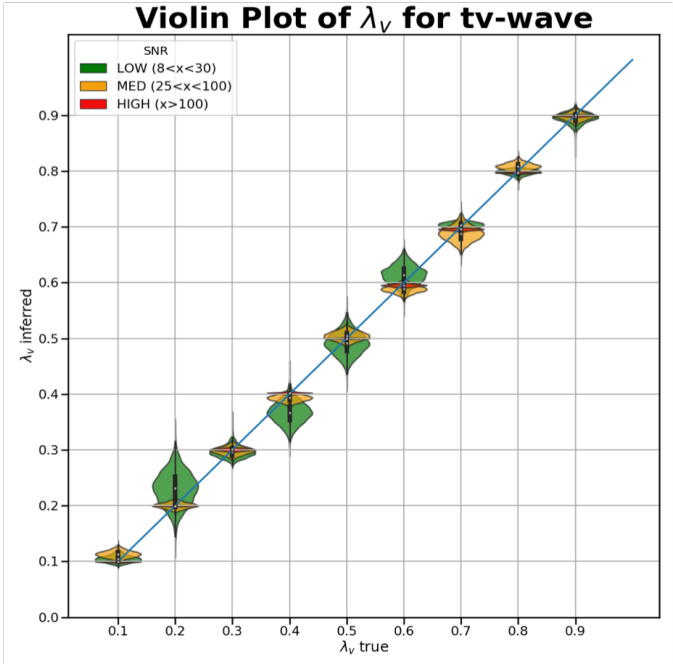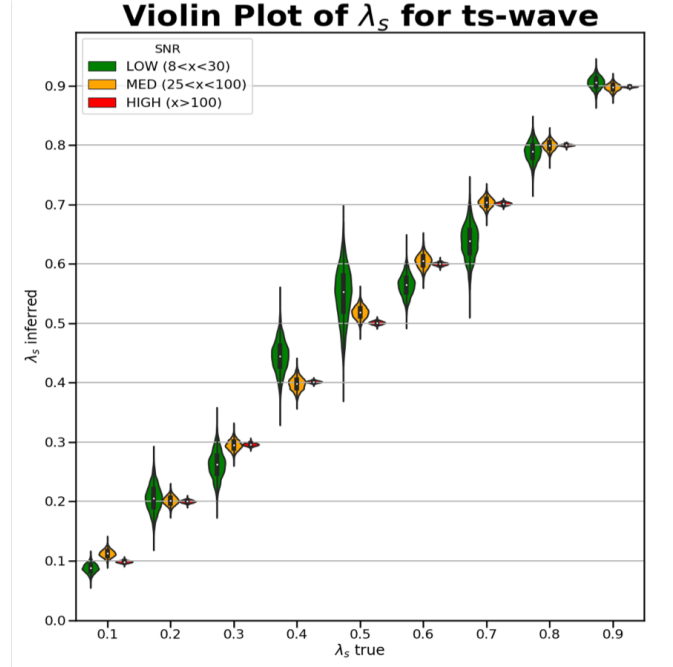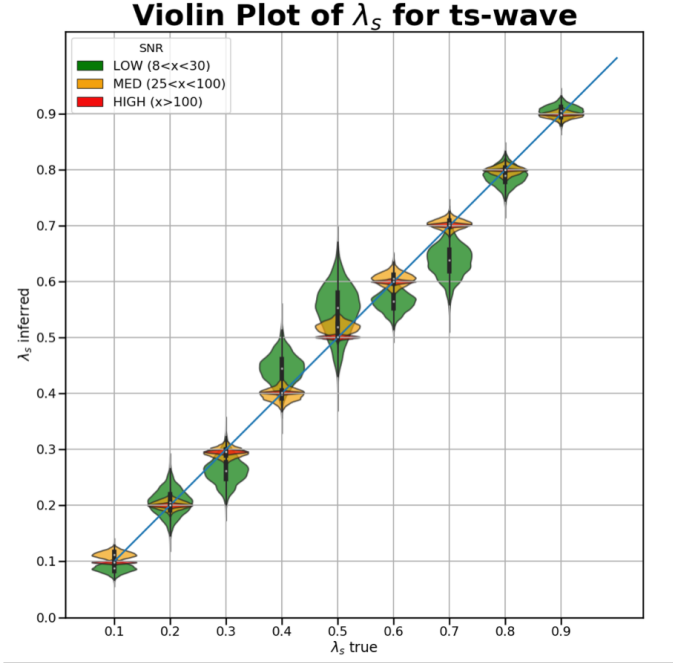

FIG. 13. *Trace plots and corner plots of a low SNR source.* The injected true value (red lines in both plots) is still well recovered after some iterations of the sampling algorithm, but with a lower precision with respect to the previous case. This specific source was randomly chosen with an ϵ_y close to zero: the algorithm correctly provides upper limits for its value.

FIG. 14. *Kernel Density Estimation (KDE, see [15]) of marginalized posterior distribution of λ_s for tensor-scalar polarized waves (top) and λ_v for tensor-vector polarized waves (bottom), with different SNR and increasing non-GR content.* Both plots for each kind of two-polarization wave represent the same data: on the x-axis is plotted the true injected value, while on the y-axis the posterior distribution as recovered by the Bayesian analysis. The difference between the two plots is only in the way results are plotted: on the left, violin plots for the three SNR sources are superimposed to show if there are systematic deviations from the line at 45°, which represent the true injected value. On the right, the three violin plots are plotted next to each other, so the horizontal grid line is the only meaningful reference for the true value. The white dot inside the distribution corresponds to the mean value, while the black band represents the first quartile interval.
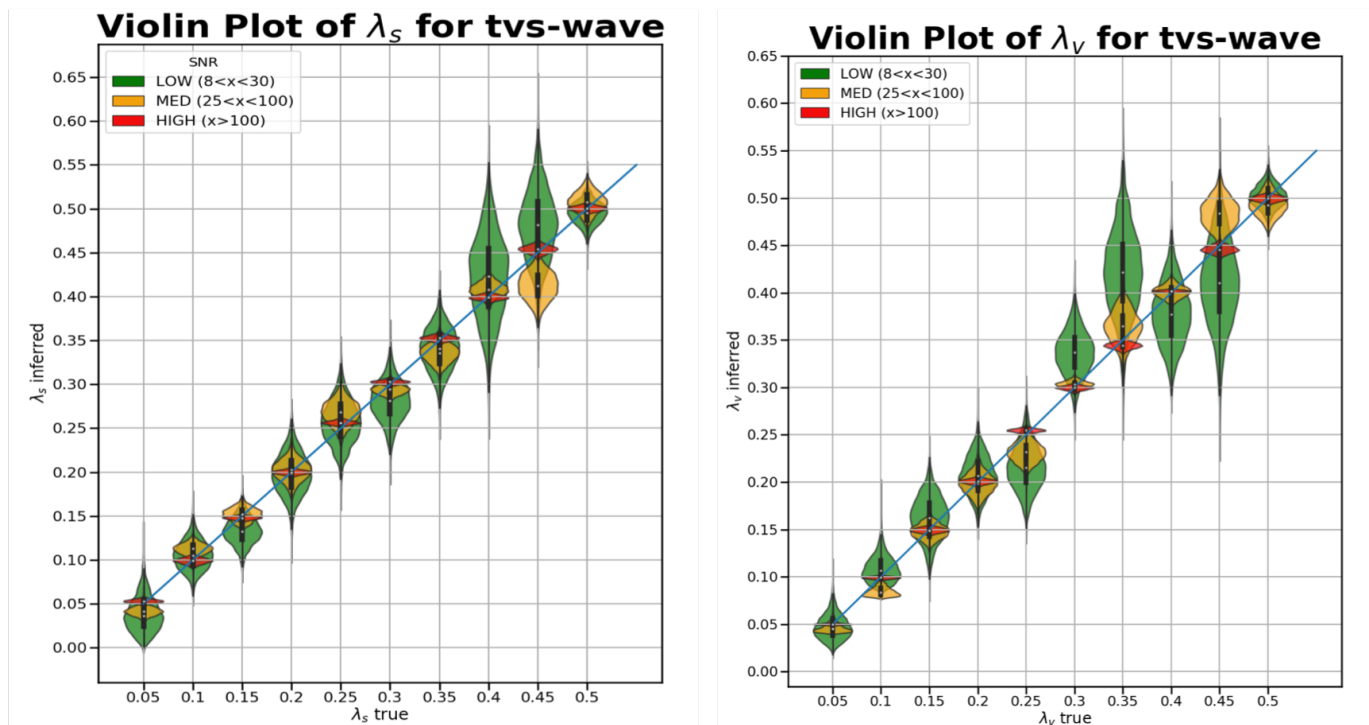
FIG. 15. *KDE of marginalized posterior distributions of λ_s (on the left) and λ_v (on the right) for tensor-vector-scalar polarized waves, with different SNR and increasing non-GR content.* On the x-axis is reported the injected true value (the same for both the lambdas), while on the y-axis the recovered distributions are superimposed for the three SNR sources. Qualitatively, the true value for low SNR is recovered generally worse in this case than the previous ones, due to the fact that a highly mixed wave generally increases the complexity and the possible degeneracies in the parameters space. The white dot inside the distribution corresponds to the mean value, while the black band represents the first quartile interval.

terior distributions for the lambdas are reported in fig.14 for the two-polarization mixture and fig.15 for the three-polarization one. As expected, the posterior distribution is narrower and well-peaked around the injected value the higher the SNR is. Qualitatively - for low SNR sources - the more the wave is a mixture of different polarizations, the worse is generally recovered by the network. This can be explained thinking that for highly mixed waves we are adding possible degeneracies to the network responses: for a complex wave (superposition of many different polarizations), it is more difficult to disentangle the single contributions. Besides this qualitative behaviour, there isn't any particular systematic on the value at which the distribution is peaked at: they are - at least for the low SNR scenario, in which they differ from the true value - randomly distributed around the true injected value.

Finally, we simulated many other sources, fixing the mean SNR (within a range of two units) for the network, in the case of a tensor-vector-scalar polarized wave with increasing non-GR content. Posterior distribution for the λ_s parameter are reported in fig.16. The disentangling power of the network generally increases with the mean SNR, with some exceptions where either the SNRs are distributed very unevenly (with a high standard deviation) between the detectors or because of statistical fluctuations in the nested sampling algorithm. This last

hypothesis is confirmed by repeating the analysis for the same data-sets multiple times, obtaining slightly different posterior distributions, compatible with the observed inversions.

VI. CONCLUSION

Since the response of a network of detectors of GWs is strictly correlated to the polarization of the wave through *antenna patterns* of each interferometer, which depend only on the relative geometry between source and detector, an extensive study of the performance of 5 detector network configurations through simulations (within a Bayesian approach) is essential.

We have found that, even using a monochromatic sine-gaussian as toy model for the waveform, the network of five L-shaped interferometers at designed sensitivity (LIGO Hanford (USA), LIGO Livingston (USA), Virgo (Italy), KAGRA (Japan) and LIGO-India) that will be available in the near future is able to efficiently disentangle the five polarization components for a generic transient signal: the precision strictly depends on the SNR of the signal and how it is distributed between the five detectors. The results were obtained within a Bayesian framework, using parameters estimation techniques.
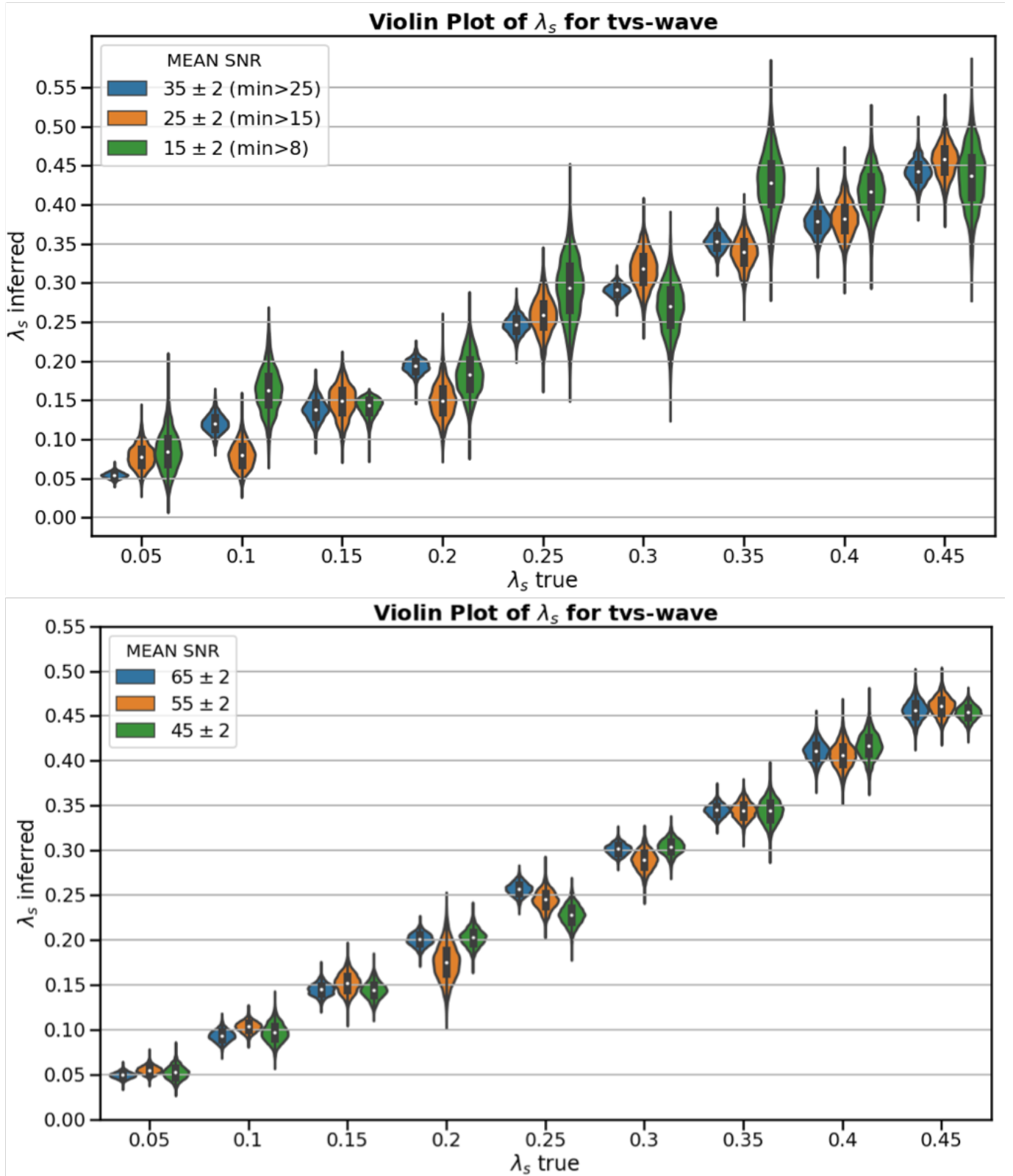
FIG. 16. *KDE of the marginalized posterior distributions of λ_s for tensor-vector-scalar polarized waves, with increasing SNR mean and non-GR content.* On the x-axis is plotted the true injected value, while on the y-axis the posterior distribution as recovered by the Bayesian analysis. The three violin plots are plotted next to each other, so the horizontal grid line is the only meaningful reference for the true value. The white dot inside the distribution corresponds to the mean value, while the black band represents the first quartile interval. Inversions in the behaviour of the recovered distribution for increasing SNR mean are due both to statistical fluctuations in the sampling algorithm and very unevenly distribution of the SNR between the five detectors.

As a first suggested future work, model selection can be used to discriminate between different models, measuring the odds of a generic theory of gravity, given the simulated data. It should be noticed that the odds as defined in Eq. (31) include by definition an Occam's penalty for models with an high number of parameters. The evidence is indeed an integral of the likelihood over all parameters space weighted by the prior function: more dimensions we add to this space (i.e. increasing the complexity of our model), bigger is the numeric value of the integral, producing smaller odds in favour of the more complex model.

Since the GW polarization is mainly a geometric factor, the results found are essentially independent of any theory model. However, it would be interesting to repeat the analysis using templates of waves from a generic metric theory of gravity, other than GR. Some of the possible waveforms for the signal have already been studied in literature (see [16]): this case will be probably addressed in a future work.

So far, we dealt exclusively with transient signals. However, it should be mentioned that if we have a longer lasting signal (possibly detected with 3G detectors or LISA), the motion of the interferometer due to Earth's rotation (or the orbital motion of the detector itself) would allow us to study the evolution in time of antenna patterns, extracting more information from the single detector. For continuous GW signals, a complete study with three ground interferometers was done in [9]. An extension of this methodology can be applied also to a larger network.

We recall that, the detection of GW non-tensor polarizations would be the first, direct evidence of new physics; at the same time, repeated non-detections would allow us to place more and more stringent tests on GR.

ACKNOWLEDGMENTS

Appendix A: One-dimensional and two-dimensional study of the likelihood

To better understand the complexity of the likelihood, a graphic study of the function in Eq. 36 for a high SNR source with random parameters is reported in the following.
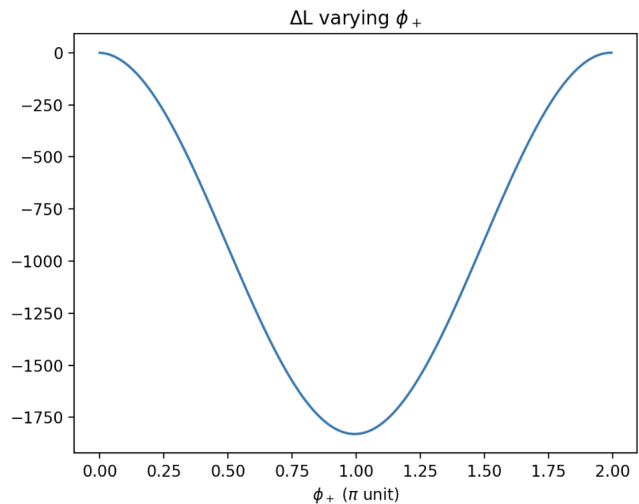

FIG. 17. *Likelihood as a function of ϕ_+.* On the y-axis the difference between the value of the likelihood at the x-value and the value of the likelihood at the injected value is plotted, while on the x-axis $\phi_+ - \phi_{true}$ in unit of π is reported.
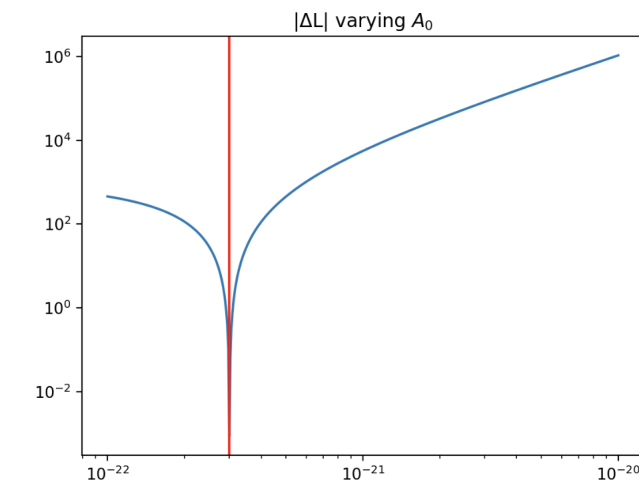

FIG. 18. *Likelihood as a function of A_0.* On the y-axis it is plotted the absolute value of the difference between the value of the likelihood at the x-value and the value of the likelihood at the injected value. On the x-axis, values for A_0 log-uniformly sampled around the injected value. The red line shows the true injected value for A_0.

In fig.17, we can see the behaviour of the likelihood as a function of one of the phases: we uniformly sampled
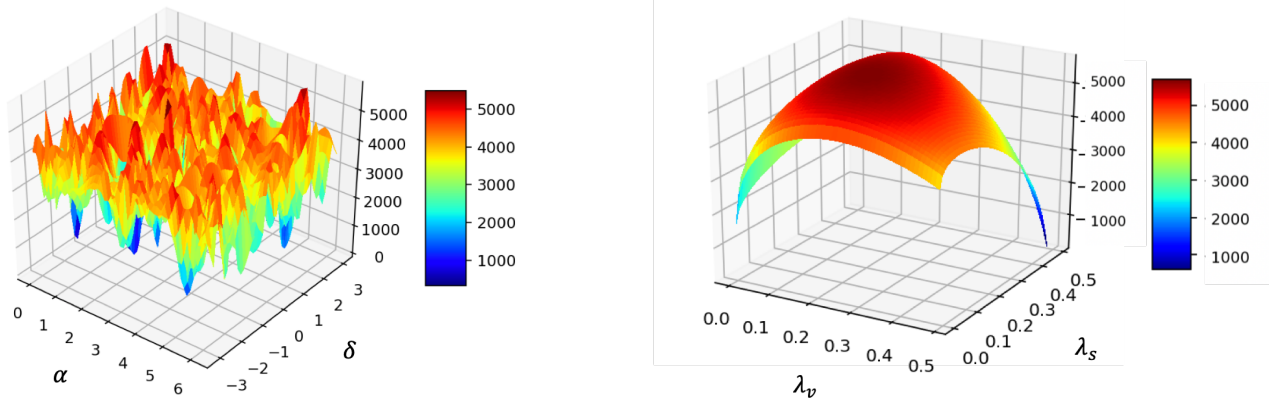
FIG. 19. *Likelihood as a function of (α, δ) (left) and (λ_s, λ_v) (right).*

a one-dimensional slice of the parameter space, allowing only one of the amplitude phases to vary (ϕ_+), starting from the injected true value. On the y-axis the difference between the value of the likelihood at the x-value minus the value of the likelihood at the injected value is plotted. If the true injected value is a global maximum for the likelihood, this difference should be always non positive and equal to zero only for the true value. We see that the plot shows the expected behaviour.

In fig.18, the likelihood as a function of the overall amplitude is studied. On the y-axis, the modulus of the difference between the value of the likelihood at the x-value and the value of the likelihood at the injected value is plotted. If the true injected value is a global maximum for the likelihood, the plot should drop down to $-\infty$ for A_0 equal to the true value. Also in this case, the plot shows the expected behaviour.

In fig.19, two bi-dimensional studies of the likelihood as a surface, in the first case function of right ascension α and declination δ and in the second one function of the two lambdas hyper-parameters are shown. In the first case, the likelihood is very complicated, with several local maxima and irregularities. This is due to the *fringe effect*, relevant for the waveform used: changing the sky location shifts the sine-Gaussians in the N detectors with respect to one another. For high frequency (f_0) signals, it is sufficient a small difference in time (whenever $f_0 \cdot \Delta t = 2n\pi$, for a small integer n) to allow the shifted sine-Gaussian to well approximate the exact waveform. This makes the likelihood itself very sensitive to a small perturbation of the two sky parameters. However, there is one global maximum, although it can be barely distinguished at the center of the plot: it corresponds to the true injected values for the sky location parameters in this specific simulation. In the second plot, the surface is smooth and regular, with a global maximum for the injected true value.

[1] E. Poisson and C. Will, *Gravity: Newtonian, Post-Newtonian, Relativistic* (Cambridge University Press, 2014).

[2] D. M. Eardley, D. L. Lee, and A. P. Lightman, Gravitational-wave observations as a tool for testing relativistic gravity, Phys. Rev. D **8**, 3308 (1973).

[3] M. Isi and A. J. Weinstein, Probing gravitational wave polarizations with signals from compact binary coalescences, (2017), arXiv:1710.03794 [gr-qc].

[4] B. Sathyaprakash *et al.*, Scientific Objectives of Einstein Telescope, *Gravitational waves. Numerical relativity - data analysis. Proceedings, 9th Edoardo Amaldi Conference, Amaldi 9, and meeting, NRDA 2011, Cardiff, UK, July 10-15, 2011*, Class. Quant. Grav. **29**, 124013 (2012), [Erratum: Class. Quant. Grav.30,079501(2013)], arXiv:1206.0331 [gr-qc].

[5] https://gwic.ligo.org/3Gsubcomm/documents/3G-observatory-science-case.pdf (2019).

[6] http://www.et-gw.eu/index.php/letter-of-intent (2019).

[7] B. P. Abbott *et al.*, Multi-messenger Observations of a Binary Neutron Star Merger, Astrophys. J. **848**, L12 (2017), arXiv:1710.05833 [astro-ph.HE].

[8] K. Chatziioannou, N. Yunes, and N. Cornish, Model-Independent Test of General Relativity: An Extended post-Einsteinian Framework with Complete Polarization Content, Phys. Rev. **D86**, 022004 (2012), [Erratum: Phys. Rev.D95,no.12,129901(2017)], arXiv:1204.2585 [gr-qc].

[9] M. Isi, M. Pitkin, and A. J. Weinstein, Probing Dynamical Gravity with the Polarization of Continuous Gravitational Waves, Phys. Rev. **D96**, 042001 (2017),

arXiv:1703.07530 [gr-qc].

[10] N. J. Cornish and T. B. Littenberg, BayesWave: Bayesian Inference for Gravitational Wave Bursts and Instrument Glitches, Class. Quant. Grav. **32**, 135012 (2015), arXiv:1410.3835 [gr-qc].

[11] A. Nitz, I. Harry, D. Brown, C. M. Biwer, J. Willis, T. D. Canton, L. Pekowsky, C. Capano, T. Dent, A. R. Williamson, S. De, M. Cabero, B. Machenschalk, P. Kumar, S. Reyes, T. Massinger, D. Macleod, A. Lenon, S. Fairhurst, A. Nielsen, shasvath, F. Pannarale, S. Khan, L. Singer, dfinstad, S. Babak, H. Gabbard, M. Tápai, P. Couvares, and L. M. Zertuche, gwastro/pycbc: Pycbc v1.13.1 release (2018).

[12] J. Skilling, Nested sampling for general bayesian computation, Bayesian Anal. **1**, 833 (2006).

[13] J. S. Speagle, dynesty: A dynamic nested sampling package for estimating bayesian posteriors and evidences (2019), arXiv:1904.02180 [astro-ph.IM].

[14] J. Veitch and A. Vecchio, Bayesian coherent analysis of in-spiral gravitational wave signals with a detector network, Phys. Rev. D **81** (2010).

[15] `https://seaborn.pydata.org/generated/seaborn.violinplot.html`.

[16] K. Chatziioannou, N. Yunes, and N. Cornish, Model-independent test of general relativity: An extended post-einsteinian framework with complete polarization content, Physical Review D **86**, 10.1103/physrevd.86.022004 (2012).

[17] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Tests of General Relativity with the Binary Black Hole Signals from the LIGO-Virgo Catalog GWTC-1, (2019), arXiv:1903.04467 [gr-qc].

[18] B. P. Abbott *et al.* (LIGO Scientific, Virgo), Tests of General Relativity with GW170817, Phys. Rev. Lett. **123**, 011102 (2019), arXiv:1811.00364 [gr-qc].

[19] C. M. Will, The confrontation between general relativity and experiment, Living Reviews in Relativity **9**, 10.12942/lrr-2006-3 (2006).

[20] B. S. Sathyaprakash and B. F. Schutz, Physics, astrophysics and cosmology with gravitational waves, Living Reviews in Relativity **12**, 2 (2009).