PHYSICAL REVIEW D **95,** 082006 (2017)

Locally optimum network detectors of unmodeled gravitational wave bursts in glitch noise

Maria Principe[*] and Innocenzo M. Pinto[†]

WavesGroup, University of Sannio at Benevento,
Piazza Roma 21, 82100 Benevento, Italy, INFN, sezione di Napoli, LVC and KAGRA

(Received 13 November 2010; revised manuscript received 8 March 2016; published 19 April 2017)

The detection of unmodeled gravitational wave transients (bursts) using a network of interferometric detectors affected by non-Gaussian (glitchy) noise is studied, starting from first principles, for the simplest case where the source position on the celestial sphere is fiducially known from different (e.g., optical, radio, or neutrino) observations. Interferometer noise modelling is preliminarily discussed in an operational perspective. Two alternative, locally optimum detectors are proposed and evaluated in simulated realistic non-Gaussian (glitchy) noise, together with robust implementations which are tolerant against incomplete knowledge or fluctuations of the noise features. These detectors outperform those based on the unrealistic stationary/Gaussian noise assumption, and in principle, do not require preliminary *ad hoc* data vetoing or laundering. They are structurally akin to those adopted in current data analysis pipelines, preserving their correlation structure, and use simple memoryless nonlinearities for data preprocessing, with only a minor added computational burden.

DOI: 10.1103/PhysRevD.95.082006

I. INTRODUCTION

The first direct detection of gravitational radiation by the two LIGO detectors [1,2] marks the opening of gravitational wave (GW) astronomy.

This milestone result was largely due to the increased sensitivity and reliability of advanced LIGO [3,4], that started operation in the second half of 2015 and observed the first signal in September 2015, during its 8th engineering run. The first and second observed GWs were transient signals emitted during the final inspiral merger and ringdown phases of two black hole binaries [1,2], for which accurate numerical and analytical waveforms are available, that helped to reconstruct the source with remarkable accuracy [5–9].

More general, short transients (bursts) of gravitational radiation may be emitted by a variety of strong-gravity cosmic events, including core-collapse supernovae, gamma ray bursters, compact object mergers, and newborn black holes (see [10] for a broad overview), forming a substantial fraction of observable GWs.

For this broader class of GW bursts (henceforth GWBs), we are facing the problem of detecting transients of basically *unmodeled* shape.

As a further complication, spurious transients of instrumental and/or environmental origin, known as glitches, appear frequently in the data channel of interferometric detectors of gravitational waves. Distinguishing GWBs from spurious glitches using data from a single detector is almost hopeless. GWBs should be accordingly sought using data from several detectors (network detection).

Remarkably, the data holding the observed GW signals were clean [11] from glitches, which allowed to claim the detection with high confidence. More second generation (advanced) detectors will come into operation soon: including Advanced Virgo [12] and the advanced version of GEO600 [13] in Europe, the large scale cryogenic japanese detector KAGRA [14] will join in a few years, plans for a LIGO-sibling observatory in India (IndIGO [15]) have been officially approved, and new players (Australia, China and Russia) may join in the mid future. The foundations of network detection of unmodeled signals were laid out in a number of seminal papers.

In the era of acoustic GW detectors [16], various coincidence algorithms based on consistency tests among candidate events gathered by different detectors were envisaged [17,18]. These laid the foundations of *noncoherent* or *coincidence-based* network data analysis [19–21].

A more effective way [22] to combine data from several wide-band interferometric detectors is to construct a *single* detection statistic (*coherent* network data analysis) to be used as a likelihood ratio [23] for classical hypothesis test [24–41].

The network likelihood ratio can be easily written, assuming independent noises in the sensors [42], and is a function of the source location on the celestial sphere (aka, the direction of arrival, henceforth DOA), and the two GW polarization components. The DOA uniquely determines the arrival time of the gravitational wave transients at the various sensors, and the directional response of each. In some cases, the DOA can be fiducially estimated from different (e.g., optical, radio, or neutrino) observations of the same cosmic events (in this case one speaks of a *triggered* search/observation).

[*]principe@unisannio.it
[†]pinto@sa.infn.it

The possibility of reconstructing the waveform directly from the (noisy) data, in a network of three or more interferometers by solving an inverse problem was first discussed by Gürsel and Tinto [25]. They also pointed out the possibility of identifying the DOA by seeking the minimum of a certain linear combination of the data themselves. Such a combination (actually, $N-2$ such combinations exist in a network of N nonaligned detectors) is known as a Sagnac-mode or null-stream [26].

The Maximum Likelihood (ML) principle [23] can also be invoked, in order to estimate unmodeled GWBs, using the (unknown) waveform time samples (or projections on a suitable functional set, e.g., sine-Gaussian waveforms [38,40,43]) to parameterize the sought signals, and seeking the supremum of the conditional likelihood ratio in the related parameter space. The supremum is accordingly used as a detection statistic, and its coordinates yield an estimate of the signal possibly hidden in the data. This idea can be traced back to Flanagan and Hughes [27] in connection with the detection of unmodeled binary black hole mergers, and it is known as the standard likelihood approach in the GW data analysis literature [28,41].

In Gaussian noise, the standard likelihood approach is equivalent to solving the Gürsel-Tinto problem via a Moore-Penrose pseudo-inverse [44].

The waveform-reconstruction problem can be unfortunately ill posed, due to source-position dependent rank-deficiency of the network response matrix, as first emphasized by Rakhmanov [29]. Ill conditioning entails several pathologies: the variance of the estimated waveforms may blow up for some source locations in the sky, and produce widely different responses in nearly (but not exactly) aligned detectors (two-detector paradox [30]). Such effects can be mitigated, using classical (e.g., Tikhonov, [29]) regularization, or maximum-entropy approaches [31].

Summing up, a well developed general framework for detecting unmodeled GWBs using a network of interferometers is available, except for the underlying ubiquitous simplifying assumption that the noise corrupting the data is stationary and Gaussian. Various *ad hoc* ideas have been proposed to deal with nonideal (glitchy) noise for the important (but special) case of *strong glitches*, i.e., glitches that can be individually detected against the (fiducially stationary and Gaussian) noise floor. At the simplest level, certain data stretches can be tagged as low quality, and discarded whenever the presence of transient disturbances of instrumental/environmental origin is detected [45]. The cut threshold can be properly tuned for each detector so as to minimize the estimated background rate of transient noise glitches [46–48].

In addition, the detection statistics appropriate to stationary Gaussian noise can be redefined taking *ad hoc* into account the amount of local transient noise, so as to reduce the false alarm probability (see e.g. [41]).

More recently, Cornish and coworkers formulated the detection problem as a *ternary* hypothesis test (the three alternative hypotheses being the observation of Gaussian noise only, of a noise transient in Gaussian noise, and of a GW signal in Gaussian noise), in a Bayesian framework [43,49,50].

This bears some similarity to the abstract approach pioneered by Kadota for dealing with signals in the presence of transient noise, according to which sporadic glitches can be detected, then *estimated*, using, e.g., a maximum *a posteriori* likelihood approach, and then subtracted from the data *before* forming the likelihood ratio [51].

Available evidence suggests that in initial, as well as in advanced, interferometers glitches occur with *all* amplitudes, their rate of occurrence being higher the smaller their amplitude (see e.g. [11,45,52]).

The effect of weak glitches is more subtle. Weak glitches are *not* individually detectable against the noise floor, and cannot be removed by vetoing. They occur with time-varying rates (Cox processes [53]) and are responsible of making the residual noise distribution non-Gaussian (*heavy tailed*) and *non* stationary. Although the rate of strong glitches has been reduced in advanced interferometers, the rate of weak glitches is the same as in the initial LIGO [45].

It is therefore important to study the problem of network detection of unmodeled transients under the assumption where the noise in each detector is *essentially* heavy tailed. This was long since recognized, as witnessed by the pioneering papers by Creighton [54], and Allen *et al.* [55,56], and is the focus of the present paper. This is no nonsense, given that, e.g. as shown in particular in [57–59], heavy-tailedness spoils in a significant measure the performance of (network) detection algorithms that would be *optimal* in the Neyman-Pearson sense for stationary Gaussian noise.

This paper is accordingly focused on the detection (and reconstruction) of unmodeled GWB using several non co-located detectors, affected by an *essentially* non-Gaussian, wide-sense stationary, glitchy noise. It completes the study program initiated in [57–59]. The analysis in Sec. IV is completely new, and the discussions in Sec. II and at the beginning of Sec. VI add substantially to our previous results. Hopefully, our results will providing non-obvious insights, paving the way to deeper understanding of the subject, useful for next generation data analysis pipelines.

We introduce a non-Gaussian noise model, including the glitching component, capitalizing on Middleton's seminal results on generalized impulsive noise; we derive its statistical properties, and show how it can be tailored to model real detector noise, being dependent only on a few gross glitch parameters, that can be easily estimated from real data.

Next, we derive appropriate forms of the likelihood ratio in the so called *threshold* or *locally optimum* approximation [60], following a rigorous approach which makes *no ad hoc* assumptions, nor does it rely on heuristic data flagging and

post-processing checks, and formulate the related detection and waveform reconstruction problems in general terms.

Finally, we present the results of numerical simulations aimed at assessing the performance of the proposed locally optimum detectors in realistic simulated glitchy noise, whose parameters were fitted on purpose to the LIGO 5th Science Run (S5) data.

The paper is accordingly organized as follows. In Sec. II we discuss in some detail the various non-Gaussian noise components, and outline a simple but general, physically motivated model of the non-Gaussian glitchy component, whose statistics can be easily computed in principle.

In Sec. III we derive the likelihood ratio for a network of interferometers, under the reasonable assumption of independent noises in the instruments, and introduce the concept of locally optimum detector in (generic) non-Gaussian noise.

In Sec. IV we extend the formalism to the detection of unmodeled waveforms, for the simplest case of triggered search (known direction of arrival). Two possible approaches are worked out in detail, where the sought unmodeled signal is either estimated from the data themselves, or treated as the realization of a random process.

In Sec. V we introduce robust implementations of the proposed detectors, aimed at handling our incomplete knowledge of the glitch noise distribution, due, in particular, to its nonstationary features.

In Sec. VI we present the results of extensive numerical Monte Carlo simulations, to illustrate the performance of the proposed detectors, and discuss the simplifying assumptions made.

Conclusions follow under Sec. VII. A few formal developments ancillary to Sec. IV are collected in the Appendix.

II. INTERFEROMETER NOISE—TOWARD A PHYSICALLY DRIVEN MODEL

The noise corrupting the data of large baseline interferometric detectors of gravitational waves exhibits a number of typical features, which are common to *all* instruments presently or planned to be in operation. These features make the statistical properties of noise *markedly* different from those of an ideal Gaussian stationary process.

The key properties of observed instrumental noise can be characterized *operationally* as follows.

As a first nonideal feature, the power spectrum of raw data contains many narrow band features (see e.g [61]). Among these, it is expedient to distinguish those due to coupling to power supply conduits, which occur at integer multiples of the electric power-line frequency, from those originating from the high-Q mechanical resonances of the wire slings hanging the interferometer terminal mirrors. The former cause *substantial* deviations of the noise distribution from the Gaussian assumption. The latter, known as *violin modes*, are linear responses to stochastic excitations of the suspending wires/strips, and hence depending on whether such excitations are Gaussian (Brownian noise) or not (e.g., creep noise due to material dislocation [62]), they contribute a (narrow band) Gaussian noise component, or a non-Gaussian one [63]. Efficient algorithms have been developed for estimating and subtracting from the data all narrow band features of known origin [64–70].

After removing these features, a look at the data in the time domain reveals the presence of further nonideal features, represented by transient disturbances (glitches) of environmental and/or instrumental origin. As already mentioned, these are a key issue when it comes to detecting unmodeled GWBs.

Most of the effort has gone so far into classifying glitches, and tracing out their origin in the machines, exploiting correlations of the main strain channel with the numerous auxiliary channels. Little effort has gone so far into *modeling* the impulsive (glitchy) noise component, understanding the impact of the glitchy component on the performance of Gaussian-noise tailored detectors, and designing detectors appropriate to such noise. This paper is aimed at moving steps in this direction.

A. A Statistical/Physical Model of Glitch Noise

Glitches have been carefully studied, as a major cause of poor detection efficiency (see e.g. [71,72]). Much work has been done on identifying generic glitch features, classifying glitch waveforms [52,73–77], and tracing out their origin [78–81].

Accumulating evidence suggests that glitches arise from random transient environmental excitations hitting some noise-susceptible detector subsystems, and reaching the data channel through pathways characterized by specific canonical (impulse) linear responses [82]. This picture is corroborated by the experimental finding that most glitches fall into a *limited* set of *typical* waveforms [79,83,84]. Its conceptual foundations may be traced back to [85].

Denote the mentioned canonical responses (aka *elementary glitches*) as $w_i(t)$, $i = 1, 2, \ldots, L$. In general, any environmental disturbance (assumed for simplicity as impulsive in time and localized in space) will enter the instrument through *several* entry points, with *different* strengths and different delays. Correspondingly, in the data channel we shall observe a *superposition* of canonical responses (i.e., a *cluster* of *elementary* glitches), viz.

$$\psi(t - t_k; \vec{a}_k) = \sum_{i=1}^{L} A_i^{(k)} w_i(t - \theta_i^{(k)}),$$
$$\text{where } \theta_i^{(k)} = t_k + \tau_i^{(k)}, \tag{1}$$

where t_k is the random firing time of the primary disturbance, and $A_i^{(k)}$ and $\tau_i^{(k)}$ are the strength and the delay, whereby this latter couples to the i—th instrument's entry point. In (1) $\vec{a}_k$ represents the array of (random) parameters

$\{A_i^{(k)}, \tau_i^{(k)} | i = 1, 2, \ldots, L\}$, which determine the *shape* of the glitch cluster.

Note that the $A_i^{(k)}$ in (1) are *not* independent, being due to the *same* environmental disturbance identified by the index k; the same is true for the $\tau_i^{(k)}$, $i = 1, 2, \ldots, L$.

Whenever the $\theta_i^{(k)}$ in (1) fall in an interval shorter than the typical time-width θ_w of the individual canonical responses w_i, the individual terms in (1) can *not* be resolved by visual inspection. However Prony's algorithm 86]] (or robust variants thereof [87]), and/or Independent Component Analysis [88], may be used to single them out.

Glitch noise can be accordingly modeled, following Middleton [89], as a *generalized shot-noise*, viz.:

$$\nu_g(t) = \sum_{k=1}^{K[T]} \psi(t - t_k; \vec{a}_k), \tag{2}$$

where $K[T]$ is a random variable representing the total number of primary disturbances occurring in the analysis window Θ, whose duration is denoted by T, and the t_k form a set of random firing times. Equation (2) was introduced in [57].

Different terms in Eq. (2) are due to *independent* disturbances, so that it is reasonable to assume the t_k in (2) as being independent and identically distributed (henceforth i.i.d.). We shall denote by λ their expected number per unit time. Experimental evidence indicates that λ fluctuates in time. We shall assume such fluctuations to occur on time scales $\gg T$. Then, according to a theorem by Hurwitz and Kač [90], the total number of terms in (2) will be ruled by a nonhomogeneous Poisson process [53], with

$$\text{prob}[K(T) = Q] = \frac{(\bar{\lambda}_T T)^Q \exp(-\bar{\lambda}_T T)}{Q!}, \tag{3}$$

with $\bar{\lambda}_T$ being the (local) average of λ in the analysis window.

The statistical properties of the glitch noise in (2) depend notably on the product between $\bar{\lambda}_T$ and the typical time-width θ_ψ (assumed $< T$) of the ψ functions in (1).

Available data from the LIGO data suggest that glitches occur typically at rates which depend on their amplitudes, smaller glitches occurring more frequently than larger ones. Strong glitches, which are *detectable* against the noise floor (using, e.g., a change-detection algorithm [91,92]), occur at typical rates up to $\bar{\lambda}_T \sim 1\ \text{s}^{-1}$, with $\bar{\lambda}_T \theta_\psi < 1$.

Pruning the data from narrow band components and *strong* glitches, leaves a *residual* noise floor which has been pictorially described as an (adiabatically) *breathing* Gaussian noise [93]. This noise is *locally* Gaussian (in the operational sense of passing successfully some statistical test of Gaussianity [94]) on sufficiently *short* timescales (typically ~1 sec); its *local* variance fluctuates adiabatically on longer timescales; and its *global* 1st-order distribution is markedly heavy tailed (non-Gaussian).

This may likely be an effect of *weak* (undetectable) glitches occurring at relatively *high(er)* rates. Note that, in the limit where $\bar{\lambda}_T \theta_\psi \gg 1$, it can be shown that the distribution of (2) becomes Gaussian, by virtue of the Central Limit Theorem, irrespective of the individual glitch shapes [89]. Hence, the higher-rate weaker glitches will likely produce a globally non-Gaussian, but wide-sense (nonstationary) locally-Gaussian noise component, whose standard deviation fluctuates on the same time scales as the underlying glitch rate.

B. Glitch Noise Statistics

The characteristic functions of the glitch noise model (2) can be computed exactly up to any order [57,89]. For subsequent developments we will need only the first order one, which can be written

$$F_g(\xi, t) = \sum_{K=0}^{\infty} \text{prob}\{K[T] = K\} F_g(\xi, t|K), \tag{4}$$

where $F_g(\xi, t|K)$ is the conditional characteristic function

$$F_g(\xi, t|K) = E\{\exp[\imath\xi \sum_{k=1}^{K} \psi(t - t_k; \vec{a}_k)]\}. \tag{5}$$

The expectation $E(\cdot)$ in (5) is taken with respect to both the firing times, t_k, and the shape parameters, $\vec{a}_k$. Assuming the pertinent distributions as time invariant in Θ, and k-independent, Eqs. (5) and (4) become, respectively

$$F_g(\xi, t|K) = \mathcal{B}(\xi, t)^K, \tag{6}$$

and

$$F_g(\xi, t) = \exp[\bar{N}(\mathcal{B}(\xi, t) - 1)], \tag{7}$$

where

$$\mathcal{B}(\xi, t) = E\{\exp[\imath\xi\psi(t - t_0; \vec{a})]\} \tag{8}$$

and $\bar{N} = \bar{\lambda}_T T$ the expected number of spurious transients in Θ. From the characteristic function $F_g(\xi, t)$ it is straightforward to compute the (central) moments of the process $\nu_g(t)$,

$$\mu_g^{(Q)} = (-\imath)^Q \frac{\partial^Q F_g(\xi, t)}{\partial \xi^Q}\bigg|_{\xi=0}, \tag{9}$$

which can be used, e.g., to approximate the PDF of $\nu_g(t)$ using Edgeworth expansion [95]. The first four central moments are explicitly given by

$$\mu_g^{(1)} = E[g(t)] = \bar{N}E[\psi(t-t_0;\vec{a})],$$
$$\mu_g^{(2)} = E[g(t)^2] = \bar{N}^2E^2[\psi(t-t_0;\vec{a})] + \bar{N}E[\psi^2(t-t_0;\vec{a})],$$
$$\mu_g^{(3)} = E[g(t)^3] = \bar{N}^3E^3[\psi(t-t_0;\vec{a})] + 3\bar{N}^2E[\psi^2(t-t_0;\vec{a})]E[\psi(t-t_0;\vec{a})] + \bar{N}E[\psi^3(t-t_0;\vec{a})],$$
$$\mu_g^{(4)} = E[g(t)^4] = \bar{N}^4E^4[\psi(t-t_0;\vec{a})] + 6\bar{N}^3E^2[\psi(t-t_0;\vec{a})]E[\psi^2(t-t_0;\vec{a})] + 3\bar{N}^2E^2[\psi^2(t-t_0;\vec{a})] + 4\bar{N}^2E[\psi(t-t_0;\vec{a})]E[\psi^3(t-t_0;\vec{a})] + \bar{N}E[\psi^4(t-t_0;\vec{a})], \quad (10)$$

where, again, the expectations are taken with respect to both t_0 and $\vec{a}$.

The first order probability density function (henceforth PDF) of the process $\nu_g(t)$ can be obtained in principle from (7) by inverse Fourier transformation. The latter, however, can be exceedingly difficult to evaluate in analytic form. In practice, it is expedient to adopt a Gaussian-mixture approximant [96] to the sought noise PDF, viz.,

$$f(x) = \sum_i \eta_i \nu_G(\mu_i, \sigma_i; x), \quad (11)$$

where ν_G is a $N(\mu_i, \sigma_i)$ density, and $\sum_i \eta_i = 1$. Gaussian mixtures are effective in representing a wide class of non-Gaussian noises [97]. Efficient (ML) algorithms are available for estimating the mixture parameters from given noise samples [98,99] based both on frequentistic and Bayesian [100] arguments. Typically, only a few terms in (11) are required to reproduce generalized shot noises with high accuracy (in terms, e.g., of distributional distance) [101]. In all cases considered hereinafter, three terms were enough [58].

C. A Realistic Noise Model

Summing up, on the basis of available results, the noise corrupting the data streams of large baseline interferometric detectors of gravitational waves presently in operation can be regarded as consisting of three additive terms:

$$n(t) = n_{NB}(t) + n_g(t) + n_{\text{floor}}(t), \quad (12)$$

representing, respectively, a narrow band component (n_{NB}), a *strong* (detectable) glitchy component (n_g), and a residual component (n_{floor}) which is still *globally* non-Gaussian, while being *locally* Gaussian (with long-term fluctuating variance) on short time scales.

Remarkably, *none* of these terms is described by a Gaussian stationary distribution, although the first and last term may include such a component, accounting respectively, for the Brownian excitation of the violin modes, and the pure thermal (Johnson) stationary noise in the instrument.

While the first term in (12) can be effectively disposed *prior* to data analysis, the *non-Gaussianity* of the remaining two terms should be properly gauged, and duly taken into account when designing and evaluating GW detection algorithms.

On the other hand, most data analysis pipelines currently in use merely adopt *ad hoc* data laundering and quality-based vetoing procedures aimed at identifying and discarding glitchy data, before applying detection/estimation tools which are designed to work (and are guaranteed to be optimal) only in Gaussian stationary noise.

D. A Working Toy Model for Glitches

Building a glitch noise simulator based on (1), (2) requires identifying the elementary glitches (canonical responses) w_i, and estimating the joint probability distributions of the A_i and τ_i in (1), from real glitchy data. This is a nontrivial task, but all needed tools are available.

Luckily, such an accurate knowledge and modeling of glitch noise is *not* needed for setting up and testing the locally optimum detectors discussed in Secs. III–V. Indeed, as we shall see, in order to implement the locally optimum detectors, discussed in Secs. III–V, only the 1st order noise PDF is needed, and only a few *gross* features of this latter are relevant. These features, as already noted in [57] are almost *independent* from the details of the glitch waveforms in (2), being basically affected only by the expected distribution of the glitch amplitudes, and the rate $\bar{\lambda}$ at which glitches occur.

Hence, we shall make here a simplest modeling assumption, which is still adequate for our present purposes.

In [57–59] we accordingly used Gabor time-frequency atoms [102], better known as Sine-Gaussian (SG) waveforms in the GWB Literature, to represent the ψ functions in (2),

$$\psi(t-t_0;\vec{a}) = A_0 \sin[2\pi f_0(t-t_0) + \phi_0] \exp\left[-\frac{(t-t_0)^2}{2\sigma_t^2}\right]. \quad (13)$$

Time-frequency atoms [103] are over-complete function sets, which can be conveniently used to represent generic transient waveforms [104]. They capture the most relevant feature of generic glitches, namely their almost compact time-frequency support. The shape parameters of Gabor atoms include the glitch amplitude A_0, center frequency f_0 and time spread σ_t [105]. The distribution of these parameters was estimated in [57] from a bunch of observed glitches (the LIGO Q-pipeline triggers) of the 5th LIGO Science Run.

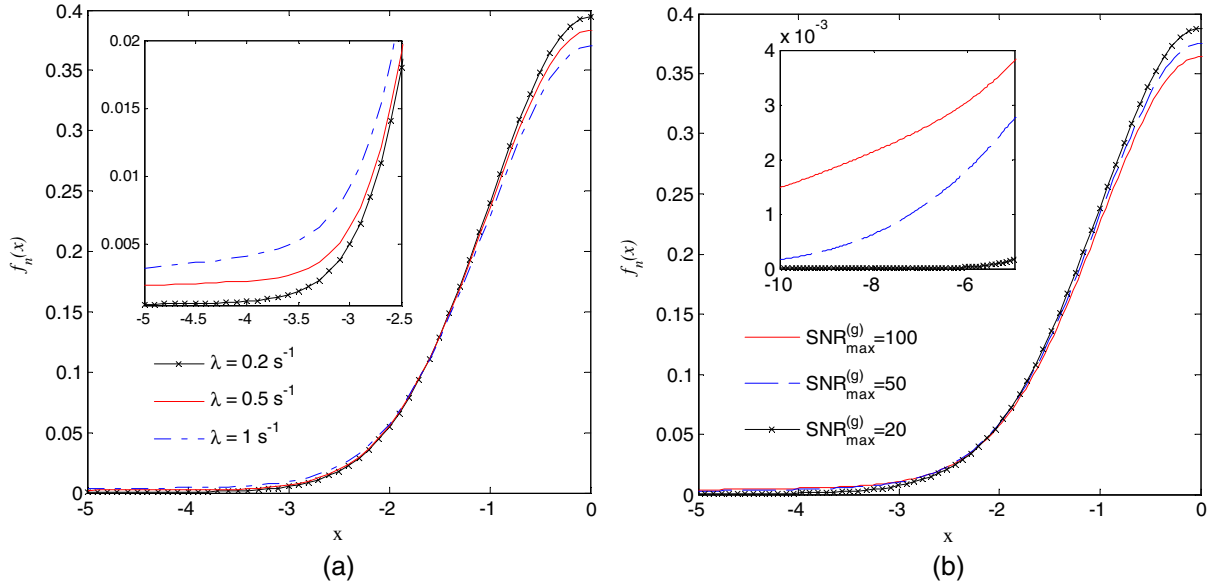

FIG. 1. First order PDF of non-Gaussian glitchy noise model, Eq. (2), superimposed to $N(0,1)$ white Gaussian noise. Glitches are modeled as in Eq. (13)—see Sec. II D for details. Top: different values of λ; $\mathrm{SNR}^{(g)}_{\mathrm{max}} = 10^2$. Bottom: different values of $\mathrm{SNR}^{(g)}_{\mathrm{max}}$; $\lambda = 0.5\ \mathrm{s}^{-1}$. Close-ups in the insets. The shown PDFs are even, and only the $x \leq 0$ range is shown.

Figure 1 shows the (1st order) PDF of the random process obtained by adding white $N(0,1)$ Gaussian noise to the random process in Eq. (2) generated using (13), for different values of the glitching rate λ and of the maximum SNR of the glitches against the Gaussian floor, henceforth denoted as $\mathrm{SNR}^{(g)}_{\mathrm{max}}$.

In Fig. 1, the glitch SNR is assumed as uniformly distributed in $[0, \mathrm{SNR}^{(g)}_{\mathrm{max}}]$, the initial phase ϕ_0 as uniformly distributed in $[0, 2\pi]$, and the distributions of f_0 and σ_t are the same as in [57].

We checked that the relevant gross features of the 1st order PDF of the random process (2) are basically the same when using either Gabor atoms (13), or the typical glitches from the database in [79], selected randomly according to their frequency of occurrence, whenever the glitch rate and SNR distribution are the same.

In Sec. VI we shall accordingly use the above toy model of glitch noise to test our proposed network detectors of unmodeled signals in a glitchy background.

III. DETECTION OF KNOWN SIGNALS WITH MULTIPLE SENSORS

Detection of known signals using multiple sensors is a well established topic in Engineering Signal Processing, originally developed in connection with radar surveillance. Two main data analysis strategies can be envisaged in this context, known as *distributed* and *centralized*. Several (e.g., tree or serial) *topologies* of distributed analysis exist (see [106] for a broad review), differing basically by the extent to which *local* decisions are taken based on the data collected at each sensor, and the way the data/decisions are eventually combined in order to construct a final decision rule. Conversely, in the centralized approach, the data gathered from all sensors are sent to a central unit which merges them to form a *single* statistic, which is used to reach a global decision.

Coincidence-based (or *noncoherent*) network data analysis methods can be regarded as a special case of distributed detection. *Coherent* methods on the other hand follow a centralized philosophy.

We consider the general case of D advanced detectors. The possibility of operating them as a single multisensor GW observatory is of paramount importance in the perspective of GW astronomy, and it is almost mandatory for detecting unmodeled GW bursts in an impulsive (glitchy) noise background.

The coherent strategy for detecting a GW signal, using a network of D interferometers, can be formulated as a binary decision problem between the alternative hypotheses H_0 and H_1 as follows,

$$\begin{cases} H_0\colon\ \mathbf{V}_d = \mathbf{n}_d \\ H_1\colon\ \mathbf{V}_d = \mathbf{S}_d + \mathbf{n}_d \end{cases}, \qquad d = 1, \ldots, D, \tag{14}$$

where boldface denotes vectors of N_s time-samples e.g. $\mathbf{V}_d = \{V_{d1}, V_{d2}, \ldots, V_{dN_s}\}$, with $V_{dk} = V_d(t_k), k = 1, \ldots, N_s$; $\mathbf{n}_d$ is the additive noise corrupting the data gathered by detector-d, and $\mathbf{S}_d$ is the GW signal received by the same detector.

For a plane gravitational wave with *transverse traceless* (TT) components $h_+(\vec{r}, t)$ and $h_\times(\vec{r}, t)$,

$$\mathbf{S}_d = F^+_d(\Omega_s)\mathbf{h}^+_d + F^\times_d(\Omega_s)\mathbf{h}^\times_d, \qquad d = 1, \ldots, D, \tag{15}$$

where $F_d^+(\Omega_s)$ and $F_d^\times(\Omega_s)$ are the *pattern functions* describing the directional response of detector-d, $\Omega_s = (\vartheta_s, \varphi_s)$ is the location of the emitting source on the celestial sphere, and $\mathbf{h}_d^{+,\times}$ are the linearly polarized TT wave components at detector-d, whose position is denoted by $\vec{r}_d$.

In the following we restrict to the case of triggered detection, where the DOA is fiducially known from different (e.g. optical, radio, or neutrino) observations [107]. It is expedient to introduce the quantity

$$h_{rss} = \left\{ \int_\Theta [h_+^2(t) + h_\times^2(t)] dt \right\}^{1/2} \tag{16}$$

and define a GW strength parameter $\theta \geq 0$, such that $\mathbf{h}_d^{+,\times} = \theta \bar{\mathbf{h}}_d^{+,\times}$, and $\bar{h}_{rss} = 1$. Also, before the subsequent processing steps, we assume that the output of detector-d is time-shifted by the (DOA dependent) delay between the wavefront arrival times at $\vec{r} = \vec{r}_d$ and $\vec{r} = 0$ (taken coincident with the Earth center), viz.,

$$\tau_d(\Omega_s) = c^{-1} \hat{n} \cdot \vec{r}_d, \tag{17}$$

where $\hat{n}$ is the unit wave-vector and c the speed of light in vacuum.

The noise process $\mathbf{n}_d$ affecting the output of detector d, is described by a generic multivariate distribution $f_{\mathbf{n}_d}(\cdot)$, representing the joint PDF of N_s noise time samples. It is reasonable to assume the noise processes in the data from different detectors as independent, in view of the large separation among the instruments. We further assume throughout this section that the noise process in any interferometer is stationary on timescales exceeding the width of the typical analysis window.

The optimum decision rule for the problem in (14) (in the Neyman-Pearson sense of yielding the minimal false-dismissal probability at a prescribed false-alarm probability) is based on computing the Likelihood Ratio (henceforth LR) test statistic [23], viz.

$$\Lambda = \frac{\prod_{d=1}^D f_{\mathbf{n}}^{(d)}(\mathbf{V}_d - \theta \bar{\mathbf{S}}_d)}{\prod_{d=1}^D f_{\mathbf{n}}^{(d)}(\mathbf{V}_d)}, \tag{18}$$

where $\bar{S}_d$ is the response (15) of detector d to $\bar{h}_d^+, \bar{h}_d^\times$, and comparing it to a false-alarm dependent threshold.

We shall henceforth assume that the noise in all interferometers has been properly whitened so that the noise time samples at the output of any detector d are i.i.d., so that [108]

$$f_{\mathbf{n}}^{(d)}(\mathbf{V}_d) = \prod_{k=1}^{N_s} f_n^{(d)}(V_{dk}), \tag{19}$$

where $f_n^{(d)}(\cdot)$ denotes the PDF of a single noise sample at the output of detector d. Taking the logarithm of the LR (18), we accordingly get the equivalent (sufficient [23]) statistic

$$\lambda' = \lg \Lambda = \sum_{d=1}^D \sum_{k=1}^{N_s} \{ \lg f_n^{(d)}[V_{dk} - \theta_0 \bar{S}_{dk}] - \lg f_n^{(d)}[V_{dk}] \}. \tag{20}$$

Note that (20) is the sum of the detection statistics of the individual detectors. For the very special case of Gaussian noise, a *uniformly* most powerful hypothesis test exists, yielding the highest detection probability at a fixed false alarm probability *irrespective* of the actual (unknown) value of $\theta > 0$, based on the well known (network) matched filter or linear correlator statistic [23],

$$\Lambda'_G = \sum_{d=1}^D \frac{\mathbf{V}_d \cdot \bar{\mathbf{S}}_d^T}{\sigma_d^2}, \tag{21}$$

where σ_d is the (Gaussian) noise standard deviation in detector-d.

It is worth noting here that the data analysis tools used in the first detections are essentially based on the following: (i) the removal of strong (detectable) glitches using data quality flags and vetoes [11,45]; (ii) an estimate of the false trigger rate using the time-slide method [109]. These are followed for unmodeled signal searches [110], by (iii) the selection of time-consistent triggers featuring excess power in the time-frequency plane (cWB [41] and oLIB [40] pipelines), followed by (iv) a waveform estimation based on the template-free constrained likelihood (cWB) or Bayesian inspired sine-Gaussian wavelet reconstruction (oLIB), and (v) the follow-up analysis (BayesWave [43,49,50]) of cWB triggers, comparing the marginalized likelihoods of three alternative hypotheses where the identified trigger data consist either of Gaussian noise, or glitch(es) in Gaussian noise, or GW in Gaussian noise. For modeled signal searches [111], triggers originate instead from (vi) template-based maximum likelihood detection and estimation (PyCBC and Gst-LAL pipelines) [111,112]. While the above framework proved to be effective in the detection of the (relatively bold) first event(s), it is not exempt from conceptual limitations, namely: (i) data vetoing trims a small yet possibly interesting fraction of the data; (ii) the time-slide method, while being free from strong modeling assumptions about the noise statistics, is poorly resilient against noise nonstationarities [113]; (iii) all pipelines use detection statistics appropriate for Gaussian and stationary noise, which is not *realistic*, modified *ad hoc* to discriminate spurious (glitch) transients. For example, PyCBC and Gst-LAL use suitable versions of the χ^2 veto introduced in [114], while the cWB uses a suitably weighted network SNR, emphasizing the

ratio between the coherent and incoherent (noise) energy in the data stream [110]; and last, (iv) the ternary hypothesis test in the Bayes Wave pipeline leaves out the possibility that a true GW signal might coexist with glitches, and hence be effectively embedded in non-Gaussian (heavy-tailed) noise.

This paper discusses network detection/estimation statistics appropriate to an *essentially* non-Gaussian background, and tests them using simulated, yet physically inspired, glitchy noise.

Several studies focused on the effect of essential (i.e., nonremovable) non-Gaussianity of the noise floor, which spoils, in a significant measure [57–59], the performance of (network) detection algorithms that would be *optimal* (in the Neyman-Pearson sense) for stationary Gaussian noise. In [115] different consistency tests augmenting matched filter data analysis, aimed at discriminating non-Gaussian noise artifacts, were tested. In [116] it was shown that performance degradation due to noise non-Gaussianity is worse in networks comprising nonaligned detectors, as, e.g., the LIGO-Virgo. In [117] it was shown that multivariate analysis (implemented via a boosted decision tree) may improve the performance of a network detector (X-pipeline) in the presence of nonideal noise features. The importance of proper assessment and modeling of non-Gaussianity of GW detector noise in the perspective of detecting low-SNR events has been recently re-emphasized in [118], where a simple (t-Student distribution based) measure of heavy tailedness was used to reveal the stationary and transient deterioration of Gaussianity in the LIGO data. Numerical simulations in [119] suggest that for suitably strong signals, parameter estimation may not be hindered by nonideal (nonstationary, non-Gaussian) noise features. However, this is certainly not the case for relatively weak (threshold) signals.

For an arbitrary non-Gaussian noise, a uniformly most powerful test does *not* exist in general. One possibility to overcome this difficulty, in the spirit of the least favorable case philosophy (see [23]), consistent with our expectations of dealing with *weak* GW signals, is to seek *weak signal* optimality, by maximizing the *slope* of the test *power* at $\theta = 0$, for a given false alarm rate. The resulting statistic and detector are called locally most powerful or *locally optimum* (henceforth LO) [120].

A. Locally Optimum Network Detector

The locally optimum network statistic for the problem (14) is given by [120]

$$\Lambda^{(\mathrm{LO})} = \sum_{k=1}^{N_s}\sum_{d=1}^{D} \frac{d}{d\theta} \lg f_n^{(d)}(V_{dk} - \theta \bar{S}_{dk})\big|_{\theta=0} = \sum_{k=1}^{N_s}\sum_{d=1}^{D} S_{dk} g_{\mathrm{LO}}^{(d)}(V_{dk}), \tag{22}$$

where

$$g_{\mathrm{LO}}^{(d)}(x) = -\frac{\partial}{\partial\theta} \lg f_n^{(d)}(\theta)\big|_{\theta=x} = -\frac{\dot{f}_n^{(d)}(x)}{f_n^{(d)}(x)} \tag{23}$$

is a nonlinear function obtained from the (first order) noise PDF. The statistic (22) can be recognized as the first coefficient in the Taylor-MacLaurin expansion of (20) in powers of θ, i.e.,

$$\lg \Lambda = \sum_{d=1}^{D}\sum_{q=1}^{\infty} \frac{(-1)^q \theta^q}{q!} \left[\sum_{k=1}^{N_s} \bar{S}_{dk}^q \frac{d^q}{dx^q} \log f_n^{(d)}(x)\big|_{x=V_{dk}}\right], \tag{24}$$

which yields a simple interpretation of $\lg\Lambda$ as a linear superposition of an infinite number of linear correlators acting on suitable memoryless nonlinear transformations of the data. This is known as a *threshold expansion* [60], where threshold refers to the limit of detectability. The locally optimum statistic is *defined* in general [120] as the lowest-order nonvanishing term in (24). The LO statistic can be cast in the following simple form

$$\Lambda^{(\mathrm{LO})} = \sum_{d=1}^{D} \mathbf{g}_d[\mathbf{V}_d] \cdot \bar{\mathbf{S}}_d^T, \tag{25}$$

where a superscript T denotes the transpose, and we used the shorthand

$$\mathbf{g}_d[\mathbf{V}_d] = \{g_{\mathrm{LO}}^{(d)}(V_{d1}), g_{\mathrm{LO}}^{(d)}(V_{d2}), \ldots, g_{\mathrm{LO}}^{(d)}(V_{dN_s})\}. \tag{26}$$

Equation (25) is clearly reminiscent of the linear correlator in Eq. (21), except that in (25) the data are preliminarily filtered by the static nonlinearity (23). Loosely speaking, the latter acts by trimming off large data samples likely due to strong glitches, which make the noise PDF heavy tailed.

For $N(0,\sigma_d)$ stationary white Gaussian noise, $g_{\mathrm{LO}}^{(d)}(x) = x/\sigma_d^2$ and Eq. (25) gives back (21).

The locally optimum detector (henceforth LOD) is obtained by comparing the LO statistic (25) to a threshold which depends only on the prescribed false alarm probability. The LOD can be proved to be *asymptotically optimal* in the limit where $\theta \to 0$ and $N_s \to \infty$ [120].

To the best of our knowledge, the LOD concept was introduced in the GW data analysis literature by Creighton [54], and discussed in considerable depth by Allen *et al.* in [55], and [56], where it was notably stressed that the LOD structure emerges naturally both in a frequentist and a Bayesian framework. We applied the LOD concept, perhaps for the first time, to the problem of detecting unmodeled GWBs in a network of interferometers affected by non-Gaussian noise [58,59].

In Fig. 2, the functions $g_{\mathrm{LO}}(x)$ obtained via (23) from the PDFs shown in Fig. 1 are displayed. It is seen that $g_{\mathrm{LO}}(x)$ is linear in an interval around $x = 0$ where the 1st order noise PDF does not depart significantly from a Gaussian distribution.

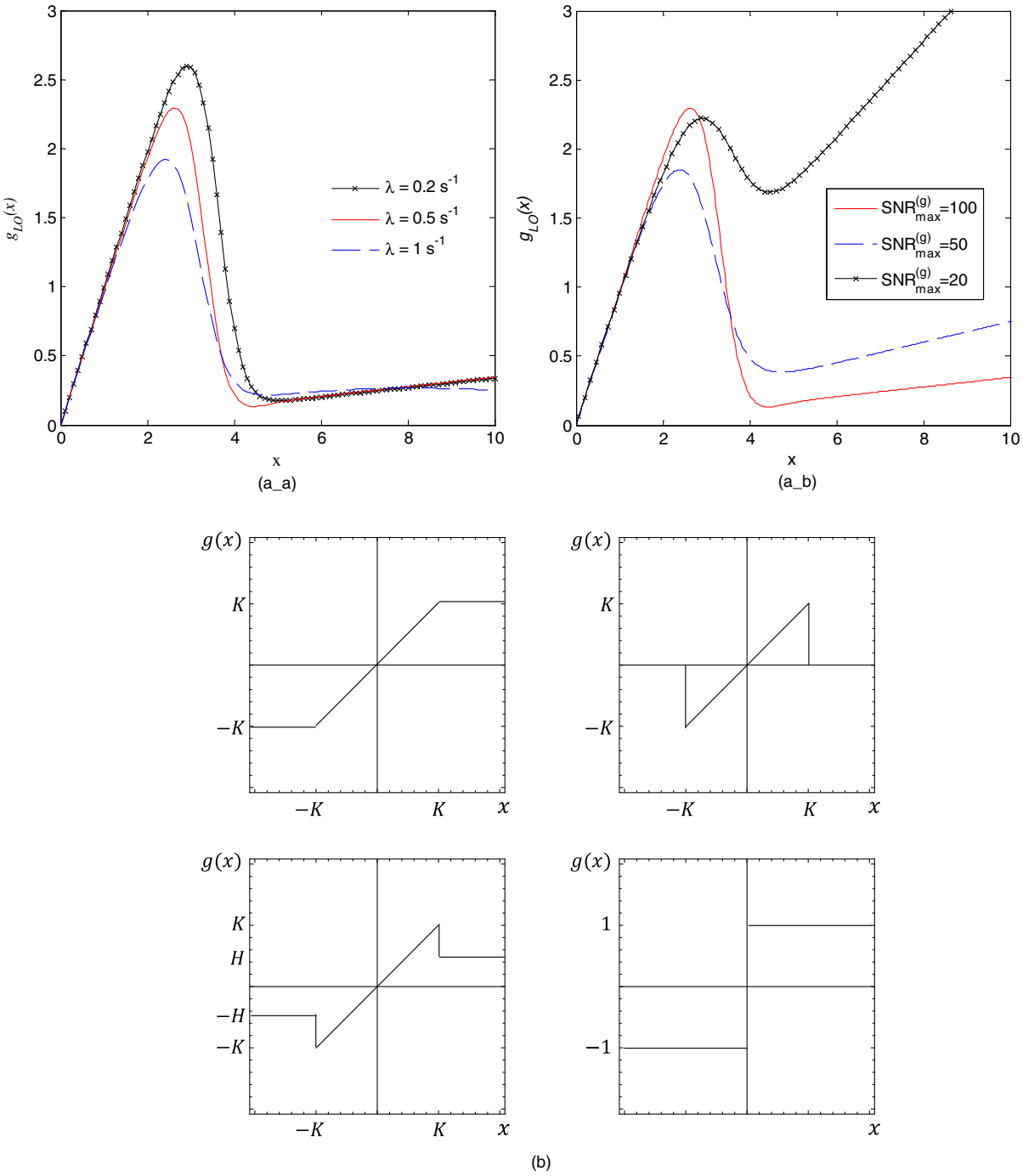

FIG. 2. (a) The function g_{LO} in Eq. (23) computed for the PDFs in Fig. 1. The displayed g_{LO} functions have odd parity, and only the $x \geq 0$ range is shown. (b) Piecewise linear approximations of the function g_{LO} in Eq. (23). Top-left: hard limiter [121]; top-right: noise blanker [122]; bottom-left: composite [123]; bottom-right: signum [124].

IV. NETWORK DETECTION OF UNMODELED GWBS

In principle, we could capitalize on *a priori* available information about the sought GWB waveforms coming from general-relativistic numerical simulations [125], to construct *templates* of the sought signals to be used in (25). However, the available numerically computed waveforms, besides being difficult to parametrize in a physically meaningful way, apply only to *a few* specific classes of sources, so that their use as templates would likely yield poor detection efficiency. Therefore, it makes more sense to assume that *no information at all* concerning the sought waveforms is available, and to seek detection schemes which may work well, irrespective of the actual shapes of

the sought signal. In this perspective, two basic approaches are possible: (a) estimating a suitable set of parameters uniquely describing the signal, possibly hidden in the data, from the data themselves; (b) looking at the sought signals as realizations of a *random* process. In the next two subsections both approaches will be discussed.

A. GWBs as Unknown Deterministic Signals—Waveform Estimation

Under the assumption of no available *a priori* information, a possible set of parameters uniquely describing the GWB signal are the very time samples of its polarization components, i.e., $\{h_k^+, h_k^\times\}, \quad \forall\ k = 1, \ldots, N_s$. This choice leads to the standard likelihood approach discussed in [28]. For an observational window including N_s time samples from D detectors, there are, accordingly, $2 \cdot N_s$ unknowns to be estimated from $D \cdot N_s$ time samples. Under the assumptions made in Sec. IV, it is readily seen that the estimation problem *breaks* up into N_s (formally identical) uncoupled problems, each yielding 2 unknowns (h_k^+ and $h_k^\times$) from D data ($V_{dk}, d = 1, 2, \ldots, D$).

The following remark is in order here. The maximum likelihood estimator is *asymptotically* optimal (in the Cramer-Rao sense [23]) when the number of data goes to infinity. In our case the number of data is D, which is only a few units. We may thus expect a poor performance of the standard likelihood estimator, irrespective of whether the noise is Gaussian or not.

On the other hand, if we had a model (even a phenomenological one) for the sought GWBs using only a *finite* number P of parameters (e.g., the principal components, as suggested in [125]), we would face the problem of retrieving P unknowns from $D \cdot N_s$ time samples [126]. The estimate would be accordingly accurate for N_s large enough, being asymptotically optimal for $N_s \to \infty$, at the expense of being most likely workable only numerically.

Letting $\mathbf{H} = \{\mathbf{h}^+, \mathbf{h}^\times\}$ the vector of unknown GW time samples (at $\vec{r} = 0$), the ML estimator maximizes the conditional PDF of the data $f(\mathbf{x}|\mathbf{H})$ over the space where $\mathbf{H}$ is supposed to lie, $\mathbf{x}$ being the actual realization of the noisy data. The coordinates, $\hat{\mathbf{H}}_{\mathrm{ML}}$, of the supremum of the likelihood ratio in the parameter space provide our estimate of the signal parameters, and the supremum itself is used as a detection statistic. In the signal processing literature this is known as the *Generalized Likelihood Ratio* (GLR) [23].

The ML estimator of $\mathbf{H}$, under the made assumption of i.i.d. noise samples, is accordingly given by

$$\begin{aligned}\hat{\mathbf{H}}_{\mathrm{ML}} &= \overset{\mathrm{argmax}}{\mathbf{H}} \prod_{d=1}^{D}\prod_{k=1}^{N_s} f_n^{(d)}(V_{dk} - S_{dk}) \\ &= -\overset{\mathrm{argmin}}{\mathbf{H}} \sum_{d=1}^{D}\sum_{k=1}^{N_s} \lg f_n^{(d)}(V_{dk} - S_{dk}),\end{aligned} \tag{27}$$

where $S_{dk} = F_d^+ h_k^+ + F_d^\times h_k^\times$. For Gaussian noise, $\hat{\mathbf{H}}_{\mathrm{ML}}$ is obtained by minimizing the following quantity

$$\mathcal{R} = \sum_{k=1}^{N_s}\sum_{d=1}^{D} \frac{(V_{dk} - S_{dk})^2}{\sigma_d^2}, \tag{28}$$

which is the sum of the σ_d^2-weighted squared residuals of the linear systems

$$\mathbf{V}_{(k)} = \mathbf{F}\mathbf{h}_{(k)}, \quad \forall\ k = 1, \ldots, N_s, \tag{29}$$

where

$$\mathbf{V}_{(k)} = (V_{1k}, V_{2k} \ldots V_{Dk})^T, \tag{30}$$

$$\mathbf{h}_{(k)} = (h_k^+, h_k^\times)^T, \tag{31}$$

and $\mathbf{F}$ is the $D \times 2$ *network response matrix*, defined as follows

$$\mathbf{F} = \begin{pmatrix} F_1^+ & F_1^\times \\ F_2^+ & F_2^\times \\ \cdots & \\ F_D^+ & F_D^\times \end{pmatrix}. \tag{32}$$

In this case, the ML estimator is the weighted least squares solution of (29), viz.

$$\hat{\mathbf{h}}_{(k)} = (\mathbf{F}^T\boldsymbol{\Sigma}\mathbf{F})^{-1}\mathbf{F}^T\boldsymbol{\Sigma}\mathbf{V}_{(k)}, \quad \forall\ k = 1, \ldots, N_s, \tag{33}$$

where $\boldsymbol{\Sigma}$ is the $D \times D$ diagonal matrix with $\Sigma_{ii} = \sigma_i^{-2}$, $i = 1, 2, \ldots, D$.

If the noise is non-Gaussian, the rhs of Eq. (27) yields a *different* (nonquadratic) measure of the residual error of (29), whose minimization can be quite complicated.

In this case the accuracy of the LS solution (33) is much too sensitive to the tail behavior in the noise PDF to even be useful as an approximate solution. Several error metrics have been proposed for inverting over-determined linear systems like (29), for the case where the noise in the data is non-Gaussian, including absolute, truncated quadratic, and bisquared error metrics [127]. A general framework for constructing such error metrics, while taking into account possible ill-conditioning of the problem, has been discussed in [128], based on the minimization of the Kullback-Leibler distance (mutual information) between the actual and estimated noise distribution. In practical cases the actual non-Gaussian noise distribution will be only loosely specified, due, e.g., to nonstationarity in time, suggesting the use of *robust* estimators [129], which are not too much sensitive to uncertainty and/or fluctuations of the noise PDF in a given class.

Hereafter, in order to have a manageable expression for the ML estimator, we shall exploit again the *weak* signal assumption. Expanding the Likelihood Function to be maximized in Eq. (27) up to the second order we obtain:

$$\begin{aligned}\lg\Lambda \approx & \sum_{k=1}^{N_s}\sum_{d=1}^{D} g_{\rm LO}^{(d)}(V_{dk})(F_d^+ h_k^+ + F_d^\times h_k^\times) \\ & + \sum_{k=1}^{N_s}\sum_{d=1}^{D}\frac{1}{2}\Gamma_{\rm LO}^{(d)}(V_{dk})[(F_d^+ h_k^+)^2 \\ & + (F_d^\times h_k^\times)^2 + 2F_d^+ F_d^\times h_k^\times h_k^+],\end{aligned} \tag{34}$$

where

$$\Gamma_{\rm LO}^{(d)}(x) = \frac{\ddot{f}_n^{(d)}(x)}{f_n^{(d)}(x)} - \left(\frac{\dot{f}_n^{(d)}(x)}{f_n^{(d)}(x)}\right)^2. \tag{35}$$

Setting the partial derivatives of the LR with respect to the unknown h_k^+ and $h_k^\times$ equal to zero, we obtain a linear system whose solutions are

$$\begin{cases} \hat{h}_k^+ = \frac{\sum_{d=1}^{D} g_{\rm LO}^{(d)}(V_{dk})(I_k^{+\times}F_d^\times - I_k^{\times\times}F_d^+)}{I_k^{++}I_k^{\times\times} - (I_k^{+\times})^2} \\ \hat{h}_k^\times = \frac{\sum_{d=1}^{D} g_{\rm LO}^{(d)}(V_{dk})(I_k^{+\times}F_d^+ - I_k^{++}F_d^\times)}{I_k^{++}I_k^{\times\times} - (I_k^{+\times})^2} \end{cases}, \qquad k = 1, 2, \ldots, N_s, \tag{36}$$

where

$$\begin{aligned} I_k^{++} &= \sum_{d=1}^{D}\Gamma_{\rm LO}^{(d)}(V_{dk})(F_d^+)^2 \\ I_k^{\times\times} &= \sum_{d=1}^{D}\Gamma_{\rm LO}^{(d)}(V_{dk})(F_d^\times)^2 \\ I_k^{+\times} &= \sum_{d=1}^{D}\Gamma_{\rm LO}^{(d)}(V_{dk})F_d^+ F_d^\times \end{aligned} \tag{37}$$

The Eqs. (36) are *Locally Optimum Estimators* (henceforth LOE) of h_k^+, $h_k^\times$. It can be shown that if the number of data D in (36) is sufficiently large, and the GW amplitudes $h_+, h_\times$ are sufficiently small ($\mathcal{O}(D^{-1/2})$), the LOE is asymptotically normal and efficient, in the Cramer-Rao sense [130–132].

The Eqs. (36) can be written as follows

$$\hat{\mathbf{h}}_{(k)} = (\mathbf{F}^T\boldsymbol{\Sigma}_{(k)}\mathbf{F})^{-1}\mathbf{F}^T\mathbf{g}_{(k)} k = 1, \ldots, N_s, \tag{38}$$

where

$$\mathbf{g}_{(k)} = (g_{\rm LO}^{(1)}(V_{1k}), g_{\rm LO}^{(2)}(V_{2k}), \ldots, g_{\rm LO}^{(D)}(V_{Dk}))^T, \tag{39}$$

and $\boldsymbol{\Sigma}_{(k)}$ is a diagonal $D \times D$ matrix, whose nonzero elements are $\Sigma_{(k)}^{dd} = -\Gamma_{\rm LO}^{(d)}(V_{dk})$. Equation (38) shares the same structure as Eq. (33), derived in the Gaussian noise case, except that here the matrix $\boldsymbol{\Sigma}_{(k)}$ depends on the time sample k, and the data are preliminary filtered through the $g_{\rm LO}^{(d)}(\cdot)$ functions. Whenever the matrices $\boldsymbol{\Sigma}_{(k)}$ are negative semi-definite, the estimator in Eq. (38) minimizes a *generalized* weighted squared residual.

For Gaussian noise $g_{\rm LO}^{(d)}(x) = x/\sigma_d^2$, and $\Gamma^{(d)}(x) = -\sigma_d^{-2}$, and Eq. (38) gives back Eq. (33). Also, Eq. (33) reduces to the Moore-Penrose pseudo-inverse based solution [44], viz.

$$\hat{\mathbf{h}}_{(k)} = (\bar{\mathbf{F}}^T\bar{\mathbf{F}})^{-1}\bar{\mathbf{F}}^T\bar{\mathbf{V}}_{(k)}, \quad \forall\ k = 1, \ldots, N_s, \tag{40}$$

where

$$\bar{\mathbf{V}}_{(k)} = (V_{1k}/\sigma_1, V_{2k}/\sigma_2 \ldots V_{Dk}/\sigma_D)^T, \tag{41}$$

and

$$\bar{\mathbf{F}} = \begin{pmatrix} F_1^+/\sigma_1 & F_1^\times/\sigma_1 \\ F_2^+/\sigma_2 & F_2^\times/\sigma_2 \\ \cdots & \\ F_D^+/\sigma_D & F_D^\times/\sigma_D \end{pmatrix} \tag{42}$$

are the noise-weighted counterparts of Eqs. (30) and (32).

1. Regularization

The network response matrix can be ill conditioned in some regions of the celestial sphere [29], and the variance of the corresponding estimator of $\mathbf{h}$ can accordingly blow up dramatically. Given a general linear system

$$\mathbf{b} = \boldsymbol{\Xi}\mathbf{h}, \tag{43}$$

it is well known that the error $\delta\hat{\mathbf{h}}$ in the solution $\hat{\mathbf{h}}$ can be fairly larger than the error $\delta\mathbf{b}$ in the data $\mathbf{b}$. The following upper bound exists

$$\frac{\|\delta\hat{\mathbf{h}}\|}{\|\hat{\mathbf{h}}^{(0)}\|} \leq \mathrm{cond}[\boldsymbol{\Xi}]\frac{\|\delta\mathbf{b}\|}{\|\mathbf{b}^{(0)}\|}, \tag{44}$$

where $\mathbf{b}^{(0)}$ and $\hat{\mathbf{h}}^{(0)}$ are the noise-free data and solution, and $\mathrm{cond}[\boldsymbol{\Xi}] = \|\boldsymbol{\Xi}\|\|\boldsymbol{\Xi}^{-1}\|$ is the condition number of matrix $\boldsymbol{\Xi}$, given by the ratio between its largest and smallest (absolute) eigenvalues [133,134]. Thus, even for mild ill conditioning (condition number $\sim$10), waveform reconstruction via (33) can be badly inaccurate [135].

Several approaches have been proposed to mitigate this problem [28,31]. In the simple regularization scheme, á la Tikhonov, one minimizes the following quantity

$$\mathcal{R} + \gamma\mathbf{h}\boldsymbol{\Omega}\mathbf{h}^T, \tag{45}$$

where $\mathcal{R}$ is the (generalized) residual of the linear system in Eq. (43), and $\boldsymbol{\Omega}$ and γ are the so called regulator matrix and

intensity, respectively. This approach results in adopting a *regularized* pseudo-inverse, such that

$$(\boldsymbol{\Xi})^{-1}_{\mathrm{reg}} = (\boldsymbol{\Xi}^T\boldsymbol{\Sigma}\boldsymbol{\Xi} + \gamma\boldsymbol{\Omega})^{-1}, \tag{46}$$

where $\boldsymbol{\Sigma}$ is the diagonal matrix of the nonunitary weights of the squared residuals. In its eigenvector space $\boldsymbol{\Xi}^T\boldsymbol{\Sigma}\boldsymbol{\Xi}$ reads

$$\begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}, \tag{47}$$

where $\mu_{1,2}$ are the eigenvalues [136], with $\mu_1 > \mu_2$. Ill conditioning occurs when $|\mu_2| \sim 0$. Rakhmanov in [29] proposed using a regulator matrix, which, in the same vector space, can be written as

$$\begin{pmatrix} 0 & 0 \\ 0 & (\mu_2\mu_1)^{1/2} - \mu_1 \end{pmatrix} \tag{48}$$

and showed that a judicious choice of $\gamma \in [0, 1]$ can make the condition number of the rhs of (46) as close to unity as desired. Obviously, some price is paid, represented by a *bias* in the waveform estimate (see [29] for a discussion).

Using the regulator (46) in (38), and the resulting estimate to compute the detection statistic (25), we obtain the locally optimum (LO) form of the generalized likelihood ratio, GLR, viz.

$$\mathrm{GLR}^{(\mathrm{LO})} = \sum_{k=1}^{N_s} \sum_{d,d'}^{1\ldots D} p_k^{dd'} g_{\mathrm{LO}}^{(d)}(V_{dk}) g_{\mathrm{LO}}^{(d')}(V_{d'k}), \tag{49}$$

where $p_k^{dd'}$ is the (d, d')-element of the $D \times D$ matrix $\mathbf{P}_k = \mathbf{F}(\boldsymbol{\Xi})^{-1}_{\mathrm{reg}}\mathbf{F}^T$.

The generalized likelihood ratio for Gaussian noise, denoted as GLR_G, is obtained by using the regulator (46) in Eq. (33), and substituting the resulting estimate in the detection statistic (21).

B. GWBs as Random Signals

The complications affecting the estimation process discussed in Subsection IV A, can be avoided if the sought signals are regarded as *realizations* of a *random* process. Adopting a maximum uncertainty attitude, we may consider $h_0^+(t)$ and $h_0^\times(t)$ as independent and identically distributed random processes, with zero mean and impulsive autocorrelation, so that

$$\begin{aligned} &E[\bar{S}_{dk}] = 0, \quad \forall\, d = 1, \ldots, D \forall\, k = 1, \ldots, N_s \\ &E[\bar{S}_{dk}\bar{S}_{pm}] = \mathcal{R}_{dp}\delta_{k-m}, \quad \forall\, d, p = 1, \ldots, D \quad \forall\, k, m = 1, \ldots, N_s, \end{aligned} \tag{50}$$

where

$$\mathcal{R}_{dp} \propto (F_d^+ F_p^+ + F_d^\times F_p^\times), \tag{51}$$

up to an irrelevant multiplying factor. Under these assumptions, the LO form of the LR statistic first derived by Kassam [137] becomes (see Appendix for details)

$$\begin{aligned} \Lambda^{(\mathrm{LO})} = &\sum_{d=1}^{D} \sum_{k=1}^{N_s} \frac{\ddot{f}_n^{(d)}(V_{dk})}{f_n^{(d)}(V_{dk})} \\ &+ 2\sum_{d=1}^{D} \sum_{p=d+1}^{D} \mathcal{R}_{dp}\mathbf{g}_d[\mathbf{V}_d] \cdot \mathbf{g}_p[\mathbf{V}_p]^T. \end{aligned} \tag{52}$$

The first term in (52) can be large even if the data consist *only* of spurious transients in the output of *individual* interferometers. The second term, known as *Generalized Cross-Correlation* (henceforth GCC), on the other hand, is a measure of the *correlations* between the outputs of *different* detectors in the network, and is nonzero only if the data contain signals of the same astrophysical origin. Thus, following Kassam, we shall drop the first term in (52) and keep only the second to form the detection statistic,

$$\mathrm{GCC} = \sum_{d=1}^{D} \sum_{p=d+1}^{D} \mathcal{R}_{dp}\mathbf{g}_d[\mathbf{V}_d] \cdot \mathbf{g}_p[\mathbf{V}_p]^T. \tag{53}$$

The structure of (53) is similar to that of (49), differing from this latter due to the absence of the diagonal terms. Interestingly, the $\mathcal{R}_{dp}$ coefficients in (53) display a very similar dependence on the source position (DOA), as the off-diagonal coefficients in (49), as shown in Fig. 3.

For Gaussian noise, Eq. (53) becomes a sum of linear correlators, henceforth denoted as LCC,

$$\mathrm{LCC} = \sum_{d=1}^{D} \sum_{p=d+1}^{D} \mathcal{R}_{dp} \frac{\mathbf{V}_d}{\sigma_d^2} \cdot \frac{\mathbf{V}_p}{\sigma_p^2}. \tag{54}$$

Remarkably, computing the GCC statistic requires *no* matrix inversion, and accordingly *no* ill-conditioning pathology may occur. On the other hand, the GCC provides *no* information about the *shape* of the detected signal, and as such it qualifies for *pure* detection purposes.

V. ROBUST IMPLEMENTATIONS

In Sections IV A and IV B we derived the locally optimum GLR and GCC detectors assuming the first order PDF of the noise processes affecting the output of all antennas as *perfectly known*. Unfortunately, this is not a realistic assumption. In practical cases the noise will be nonstationary in time, and its PDF will be loosely known.

Whenever the noise distributions are not precisely known, or are varying in time, a workable approach consists in identifying a functional class where the noise distributions can be assumed to lie or fluctuate, and adopting a detector which performs *well* (in a sense to

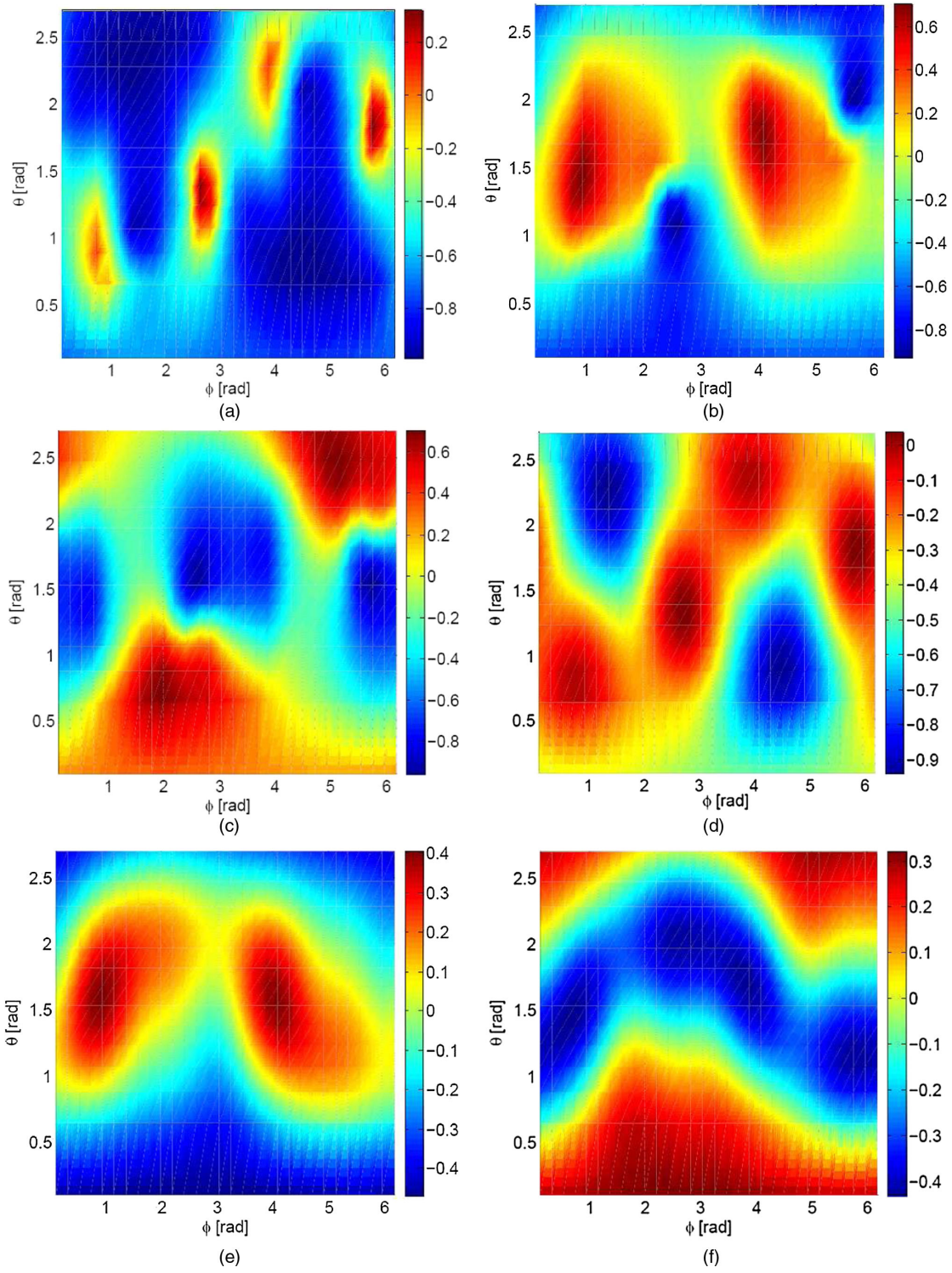

FIG. 3. Left column panels: sky maps of the coefficients $p^{dd'}$ in Eq. (49); right column panels: sky maps of the coefficients $\mathcal{R}_{dp}$ in Eq. (53). The top, middle, and bottom row panels correspond to the H1-L1, H1-Virgo, and L1-Virgo indexes, respectively.

be specified below) over the *whole* class. Clearly, the coarser the knowledge about the noise, the wider the class to which the noise supposedly belongs. We speak of robust detectors whenever the noise distribution is known to lie in a *neighborhood* of some nominal function.

We shall focus here on the class of noise distributions (henceforth denoted as $\mathcal{C}$) whose first order PDF can be written:

$$f_n(x) = (1-\epsilon)f_N(x) + \epsilon h(x), \tag{55}$$

where $f_N(x)$ is N$(0,\sigma_G)$, $h(\cdot)$ is a *generic* zero-median non-Gaussian probability density, and $\epsilon \in [0,1]$ is a mixture parameter. The noise PDF discussed in Sec. II can be recognized as a special case of (55), where

$$\epsilon = 1 - \exp(-\lambda T) \tag{56}$$

and [see Eq. (7)]

$$\epsilon h(x) = \mathcal{F}^{-1}_{\xi\to x}\{\exp(-\lambda T)[\exp(\lambda T \mathcal{B}(\xi)) - 1]\} * f_N(x), \tag{57}$$

$\mathcal{F}$ being the Fourier transform operator, and $*$ denoting convolution.

In order to implement a *robust* detector in $\mathcal{C}$, we follow the *Min-Max* strategy, i.e., we use for *all* noises with PDF in $\mathcal{C}$ the *best* (LOD) detector corresponding to the *worst* noise distribution [129]. The least favorable $\tilde{f}_n(x) \in \mathcal{C}$ is the one which *minimizes* the Fisher information [129]

$$\mathcal{I}[f_n] = \int_{-\infty}^{\infty} dx f_n(x) \left(\frac{f_n'(x)}{f_n(x)}\right)^2. \tag{58}$$

Correspondingly, the detector performance (namely, its power, at fixed false alarm level) for any other $f_n \in \mathcal{C}$ will never be worse than that for $\tilde{f}_n$. It can be shown that the least favorable distribution in $\mathcal{C}$ has the following explicit form [138]

$$\tilde{f}_n = \frac{\exp[-q(x)]}{\sigma_G\sqrt{2\pi}}, \tag{59}$$

where

$$q(x) = \begin{cases} \frac{x^2}{2\sigma_G^2}, & |x| < K \\ \frac{K|x|}{\sigma_G^2} - \frac{K^2}{2\sigma_G^2}, & |x| \geq K \end{cases}, \tag{60}$$

and thus consists of a central Gaussian bulge merging, for $|x| > K$, into exponential tails. The locally optimum nonlinearity computed from (23) using Eqs. (59), (60) is the *hard limiter* (henceforth HL) function [121],

$$g_{HL}(x) = -\frac{\dot{\tilde{f}}_n}{\tilde{f}_n} = \begin{cases} x, & |x| \leq K \\ K\mathrm{sgn}(x), & |x| > K \end{cases}, \tag{61}$$

The corresponding locally optimum coefficient (35) obtained using (59) and (60) is [139]

$$\Gamma_{\mathrm{HL}}(x) = \frac{\ddot{\tilde{f}}_n(x)}{\tilde{f}_n(x)} - \left(\frac{\dot{\tilde{f}}_n(x)}{\tilde{f}_n(x)}\right)^2 = \begin{cases} -\sigma_G^{-2}, & |x| \leq K \\ 0, & |x| > K \end{cases}. \tag{62}$$

In (61), $\mathrm{sgn}(\cdot)$ is the Dirichlet signum function, and the parameter K is related to ϵ by [138]:

$$\int_{-K}^{K} f_N(x)dx + \frac{2\sigma_G^2}{K} f_N(K) = \frac{1}{1-\epsilon}, \tag{63}$$

which follows from the obvious unit-area property of $\tilde{f}_n$.

Equations (63) and (56) can be combined to relate the parameter K in (60) to the glitch-rate λ,

$$\int_{-K}^{K} f_N(x)dx + \frac{2\sigma_G^2}{K} f_N(K) = \exp(\lambda T). \tag{64}$$

If the glitch rate fluctuates in $[0,\lambda_{\max}]$, and all other relevant noise parameters are assumed known, the least favorable distribution is obtained by using the *largest* admissible λ in (64).

The Min-Max robust implementation of the locally optimum GLR and GCC network detectors (Eqs. (49) and (53), respectively, over the broad class $\mathcal{C}$ of noises with PDF given by (55), is thus obtained using (61) in place of (23) and (62) in place of (35), with K given by (64).

We mention in passing that other piecewise-linear approximants of the LO nonlinearity have been proposed besides the HL, including the so called noise blanker [122], and composite nonlinearity [123], sketched in Fig. 2(b). All of these piecewise-linear functions are clearly reminiscent of the optimal $g_{\mathrm{LO}}(x)$ nonlinearities shown in Fig. 2(a). Heuristic prescriptions for setting the breakpoints, levels, and slopes in these approximate nonlinearities are discussed, e.g., in [123].

A rather special role is played by the signum nonlinearity, also sketched in Fig. 2(b). The signum nonlinearity is the locally-optimum one, Eq. (23), for the rather extreme case of Laplace (double exponential) distributed noise [140]. The Laplace is the least favorable distribution in the *extremely wide* class of PDFs having zero median, and the signum nonlinearity LOD is accordingly a paradigm of *nonparametric* detector [124,141].

The GCC statistic (53) using the signum nonlinearity is a *generalized polarity-coincidence* statistic [120].

The GLR statistic, on the other hand, can *not* be implemented straightforwardly for Laplace distributed noise, because the corresponding $\Gamma_{\mathrm{LO}}(x)$, Eq. (35), is identically zero [142]. Remarkably, even the simplest nonparametric signum-detector, which can be regarded as a fiducially minimum-performance detector, may outperform the linear correlator in simulated non-Gaussian noise.

VI. NUMERICAL SIMULATIONS

Before illustrating the results of numerical simulations aimed at evaluating the performance of the network detection statistics discussed in the previous section, some remarks are in order to justify the simplifying assumptions made.

The comparative performance of threshold detectors can be conveniently gauged in terms of their *asymptotic relative efficiency* (henceforth ARE), viz. [143,144]

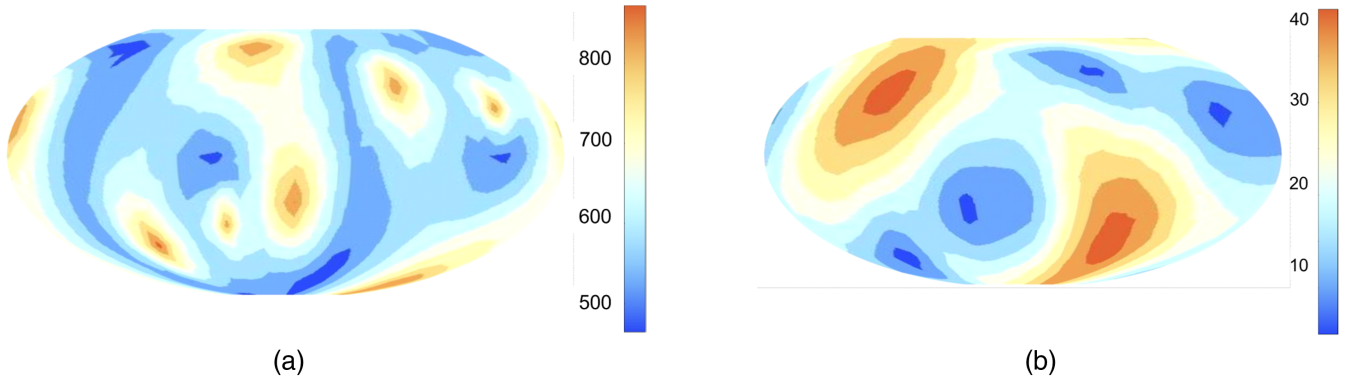

FIG. 4. Mollweide maps of detection thresholds for a false alarm probability of 10^{-2}. Non Gaussian glitchy noise with $\lambda = 0.5\ \mathrm{s}^{-1}$, $\mathrm{SNR}_{\mathrm{max}}^{(g)} = 10^2$. Top: $\mathrm{GLR}^{(\mathrm{LO})}$ detector, Eq. (49); bottom: GCC detector, Eq. (53).

$$\mathrm{ARE}(\alpha,\beta) = \lim_{\substack{N_1,N_2\to\infty \\ \theta_0\to 0}} \frac{N_2(\alpha,\beta,\theta_0)}{N_1(\alpha,\beta)}, \tag{65}$$

where N_1 is the number of data samples used by some reference algorithm (e.g., the linear correlator) to achieve a false-dismissal probability β, at a prescribed false-alarm probability α, N_2 is the number of data samples used by the detector under test, and θ_0 is the signal strength defined in Sec. III.

The ARE was introduced in [145], where it was also proved that the LOD (25) and the *strictly optimal* NP detector (20) have *equal AREs*, and its use as a synthetic detector performance index was validated by extensive simulations [146]. Under the i.i.d. assumption, the ARE with respect to the linear correlator can be computed in closed form, yielding:

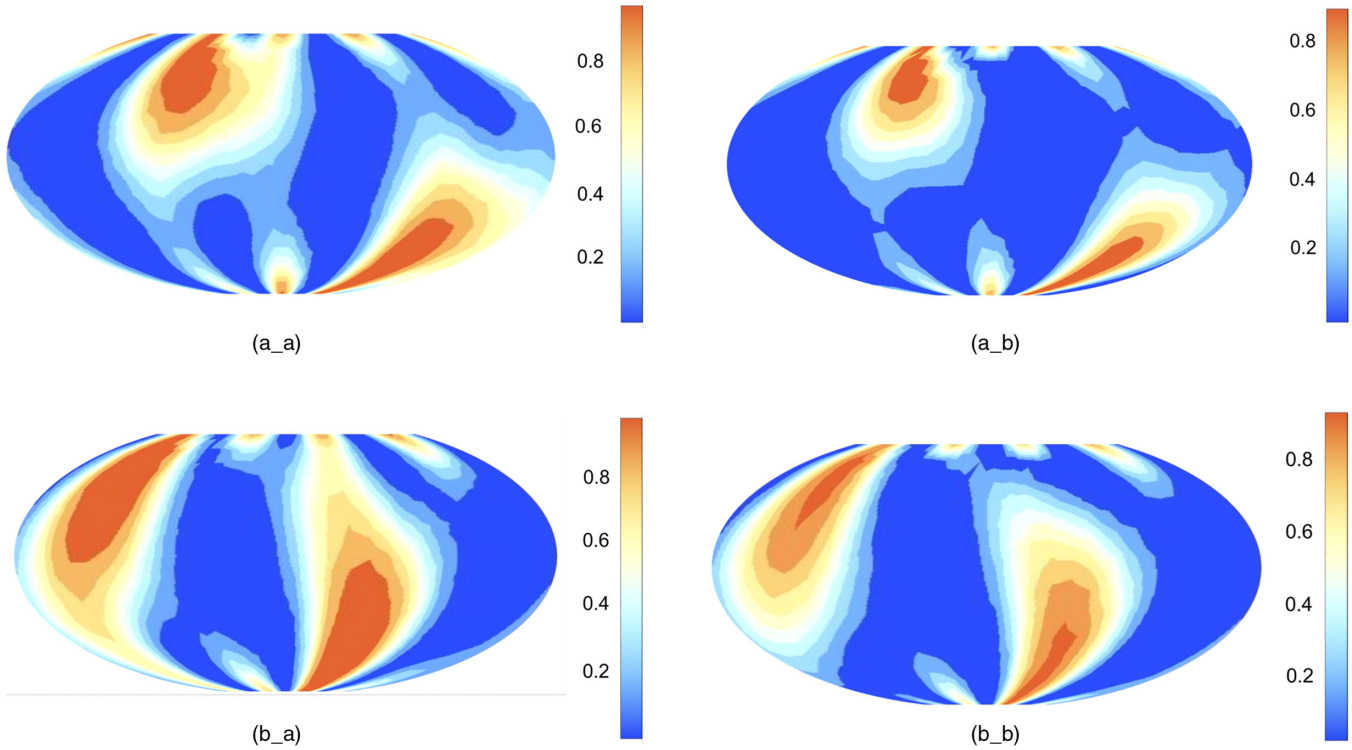

FIG. 5. (a) Mollweide maps of detection probability at a false alarm level of 10^{-2}; linearly (+)-polarized GWB with $\delta_h = 20$. Data corrupted by pure Gaussian noise. Top: Gaussian-noise version (GLR_G) of generalized likelihood ratio detector, Eq. (49); bottom: Gaussian-noise version (LCC) of generalized cross-correlation detector, Eq. (53). (b) Mollweide maps of detection probability at a false alarm level of 10^{-2}; linearly (×)-polarized GWB with $\delta_h = 20$. Data corrupted by pure Gaussian noise. Top: Gaussian-noise version (GLR_G) of generalized likelihood ratio detector, Eq. (49); bottom: Gaussian-noise version (LCC) of generalized cross-correlation detector, Eq. (53).

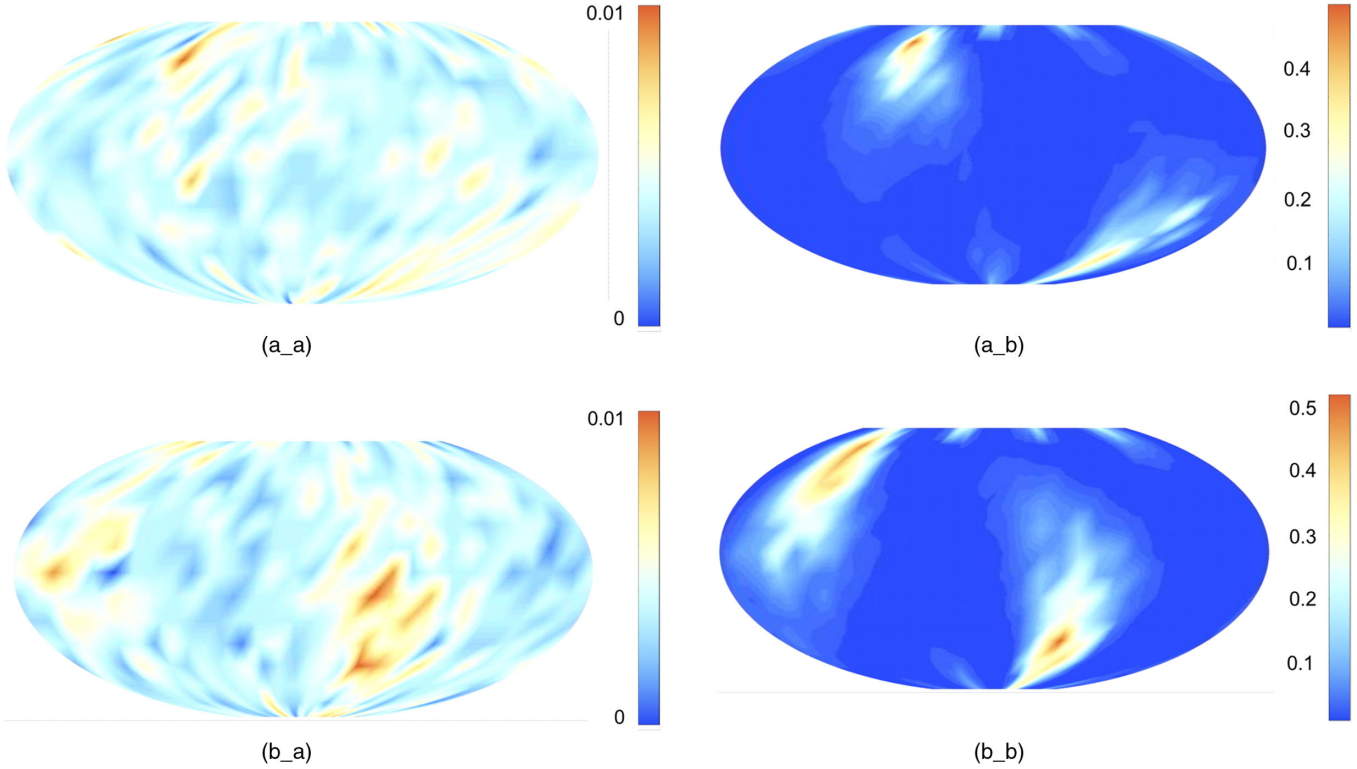

FIG. 6. (a) Mollweide maps of detection probability at a false alarm level of 10^{-2}; linearly (+)-polarized GWB with $\delta_h = 20$. Data corrupted by non-Gaussian glitchy noise, with $\lambda = 0.5\ \mathrm{s}^{-1}$, $\mathrm{SNR}^{(g)}_{\mathrm{max}} = 10^2$. Top: Gaussian-noise version (GLR_G) of generalized likelihood ratio detector, Eq. (49); bottom: Gaussian-noise version (LCC) of generalized cross-correlation detector, Eq. (53). (b) Mollweide maps of detection probability at a false alarm level of 10^{-2}; linearly (×)-polarized GWB with $\delta_h = 20$. Data corrupted by non-Gaussian glitchy noise, with $\lambda = 0.5\ \mathrm{s}^{-1}$, $\mathrm{SNR}^{(g)}_{\mathrm{max}} = 10^2$. Top: Gaussian-noise version (GLR_G) of generalized likelihood ratio detector, Eq. (49); bottom: Gaussian-noise version (LCC) of generalized cross-correlation detector, Eq. (53).

$$\mathrm{ARE} = \frac{\sigma_G^2 [\int_{-\infty}^{\infty} dx \dot{g}(x) f_n(x)]^2}{\int_{-\infty}^{\infty} dx g^2(x) f_n(x) - [\int_{-\infty}^{\infty} dx g(x) f_n(x)]^2}, \tag{66}$$

where f_n and σ_G are the first order noise PDF and the standard deviation of the Gaussian noise component in (55), respectively, and $g(x)$ is the (*not* necessarily LO) nonlinearity used to implement the threshold detector. The optimal choice $g = g_{\mathrm{LO}} = \dot{f}/f$ yields the following *upper bound* for the ARE:

$$\mathrm{ARE}_{\mathrm{max}} = \sigma_G^2 \mathcal{I}[f_n], \tag{67}$$

where $\mathcal{I}[f_n]$ is the Fisher information, Eq. (58).

A considerable body of technical literature deals with the way ARE depends on the *features* of the noise PDF, and with the (related) problem of ARE degradation due to the use of *approximate* (in particular, robust) implementations of the nonlinearity g. The main conclusions of these investigations can be summarized as follows (see, e.g. [123,143]). In the ideal case where the LOD nonlinearity is perfectly matched to the noise PDF, only the *gross* features of this latter do affect the (asymptotic) detector performance, as measured by (67). Loosely speaking, as suggested by (66), what matters here is only the *width* of the central (Gaussian) bulge of the 1st order distribution, and the *slope* of its heavy tails.

Numerical experiments based on the generalized shot noise model (2) show that these features depend essentially on the glitch rate and glitch amplitude distribution, the glitch shape being almost irrelevant.

Using approximate implementations of the LO nonlinearity also has little effect on the ARE. Miller and Thomas [123] considered several of them, including the HL, noise blanker and composite nonlinearities displayed in Fig. 2(b), and showed that they all yield comparable performances, provided their relevant parameters (slopes, breakpoints, and levels) are chosen consistently. Based on these findings, in our numerical simulations discussed below, we adopt the toy model for glitchy noise described in Sec. II D, to generate glitchy-noise instances, in order to test our LODs [147].

In our numerical experiments, we restrict to the case of three interferometers network, having in mind the present generation of advanced detectors. The noise is generated independently in each detector, and consists of the superposition of an i.i.d. $N(0,1)$ Gaussian floor, and a (pure) glitchy component generated via Eq. (2), using the Gabor

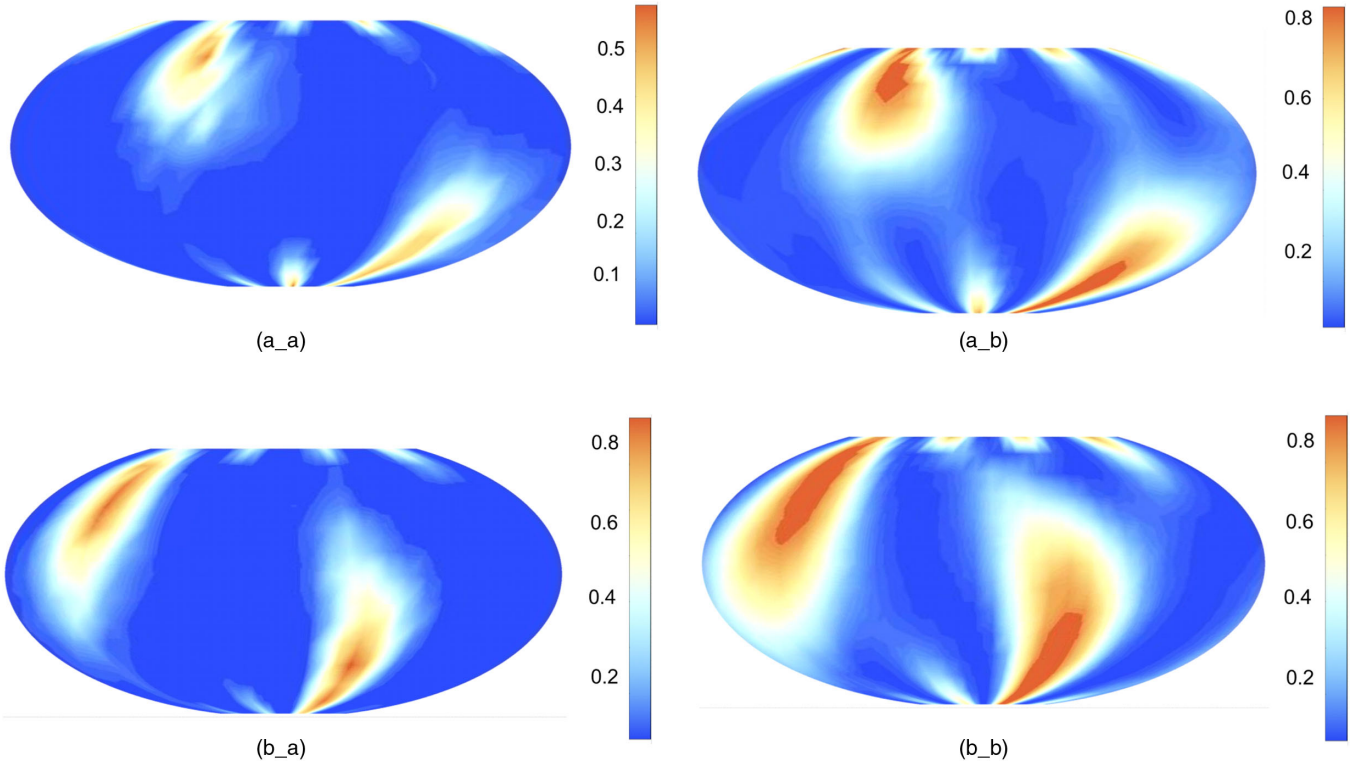

FIG. 7. (a) Mollweide maps of detection probability at a false alarm level of 10^{-2}; linearly (+)-polarized GWB with $\delta_h = 20$. Data corrupted by non-Gaussian glitchy noise, with $\lambda = 0.5\ \mathrm{s}^{-1}$, $\mathrm{SNR}^{(g)}_{\mathrm{max}} = 10^2$. Top: generalized likelihood ratio detector $\mathrm{GLR}^{(\mathrm{LO})}$, Eq. (49); bottom: generalized cross-correlation detector GCC, Eq. (53). (b) Mollweide maps of detection probability at false alarm level of 10^{-2}; linearly (×)-polarized GWB with $\delta_h = 20$. Data corrupted by non-Gaussian glitchy noise, with $\lambda = 0.5\ \mathrm{s}^{-1}$, $\mathrm{SNR}^{(g)}_{\mathrm{max}} = 10^2$. Top: generalized likelihood ratio detector $\mathrm{GLR}^{(\mathrm{LO})}$, Eq. (49); bottom: generalized cross-correlation detector GCC, Eq. (53).

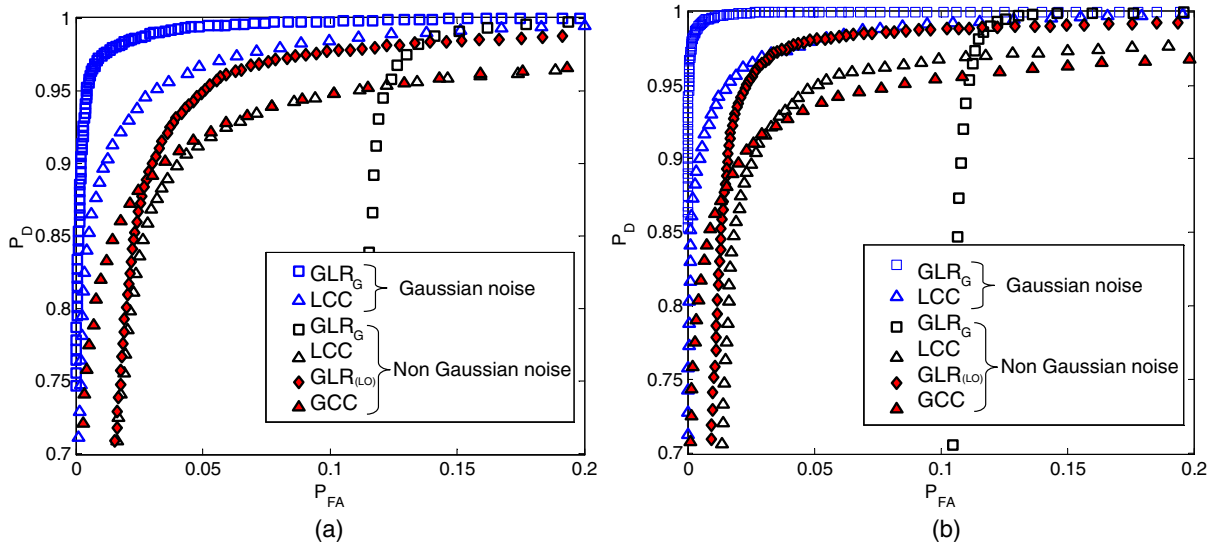

FIG. 8. (a) ROCs of the locally optimum generalized likelihood ratio detector, $\mathrm{GLR}^{(\mathrm{LO})}$, Eq. (49), and the generalized cross-correlation detector, GCC, Eq. (53), in non-Gaussian glitchy noise with $\lambda = 0.5\ \mathrm{s}^{-1}$, $\mathrm{SNR}^{(g)}_{\mathrm{max}} = 10^2$ (red markers). The ROCs of the corresponding Gaussian-noise versions, GLR_G and LCC, are also shown for comparison, both in (the same) non-Gaussian glitchy noise (black markers), and in Gaussian noise (blue markers). Linearly (+)-polarized GWB with $\delta_h = 20$; source at $\vartheta = 2.28$ rad, $\varphi = 1.99$ rad. (b) ROCs of the locally optimum generalized likelihood ratio detector, $\mathrm{GLR}^{(\mathrm{LO})}$, Eq. (49), and the generalized cross-correlation detector, GCC, Eq. (53), in non-Gaussian glitchy noise with $\lambda = 0.5\ \mathrm{s}^{-1}$, $\mathrm{SNR}^{(g)}_{\mathrm{max}} = 10^2$ (red markers). The ROCs of the corresponding Gaussian-noise versions, GLR_G and LCC, are also shown for comparison, both in (the same) non-Gaussian glitchy noise (black markers), and in Gaussian noise (blue markers). Linearly (×)-polarized GWB with $\delta_h = 20$; source at $\vartheta = 2.28$ rad, $\varphi = 0.94$ rad.

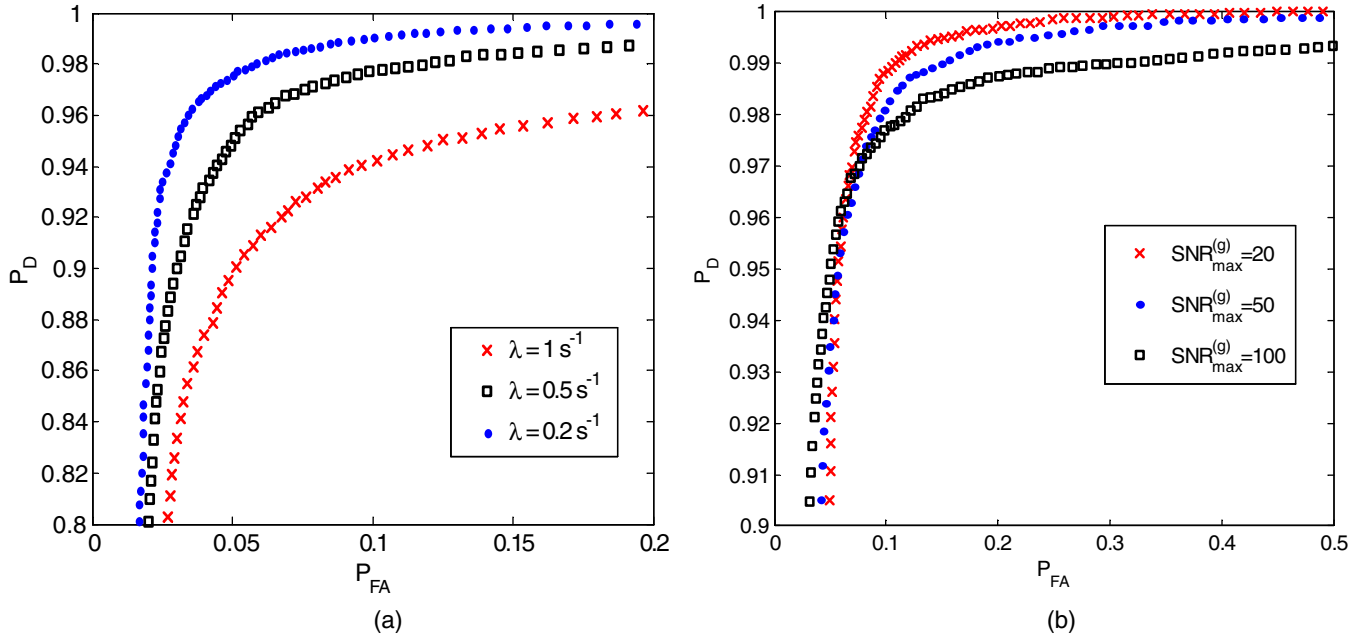

FIG. 9. ROCs of locally optimum generalized likelihood ratio detector $\mathrm{GLR}^{(\mathrm{LO})}$, Eq. (49), in non-Gaussian glitchy noise. Top: different values of λ, with $\mathrm{SNR}^{(g)}_{\mathrm{max}} = 10^2$; bottom: different values of $\mathrm{SNR}^{(g)}_{\mathrm{max}}$, with $\lambda = 0.5\ \mathrm{s}^{-1}$. Linearly (+)-polarized GWB with $\delta_h = 20$; source at $\vartheta = 2.28$ rad, $\varphi = 1.99$ rad.

(SG) atoms, Eq. (13), to mimic the glitches, whose parameter distributions are chosen as explained in Sec. II D. Obviously, the resulting random processes would *not* be i.i.d., its correlation spanning a typical number m of samples of the order of those contained in a single glitch. In this case, one technically speaks of an m-dependent process (meaning that data samples further than m samples apart are independent and dependent otherwise).

The resulting process would be similar to what would be obtained in a real experiment, if raw glitchy data were (fiducially) whitened using an estimate of the PSD of the (*plain*, colored) Gaussian background, obtained from some fiducially glitch-free data.

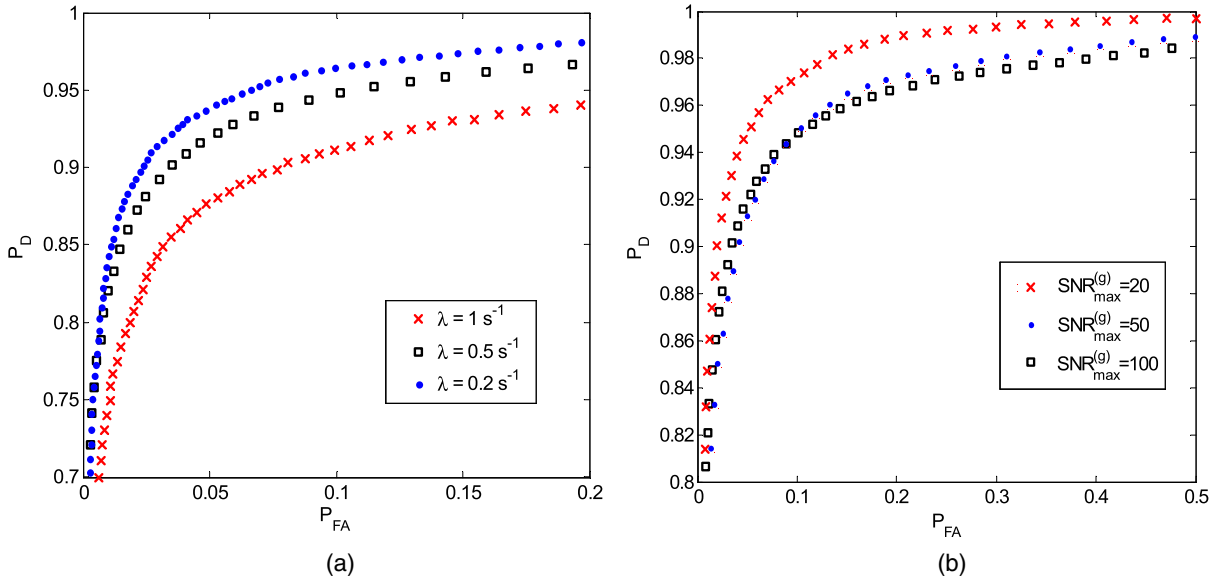

FIG. 10. ROCs of generalized cross-correlation detector GCC, Eq. (53), in non-Gaussian glitchy noise. Top: different values of λ, with $\mathrm{SNR}^{(g)}_{\mathrm{max}} = 10^2$; bottom different values of $\mathrm{SNR}^{(g)}_{\mathrm{max}}$, with $\lambda = 0.5\ \mathrm{s}^{-1}$. Linearly (+)-polarized GWB with $\delta_h = 20$, emitted by a source at $\vartheta = 2.28$ rad, $\varphi = 1.99$ rad.

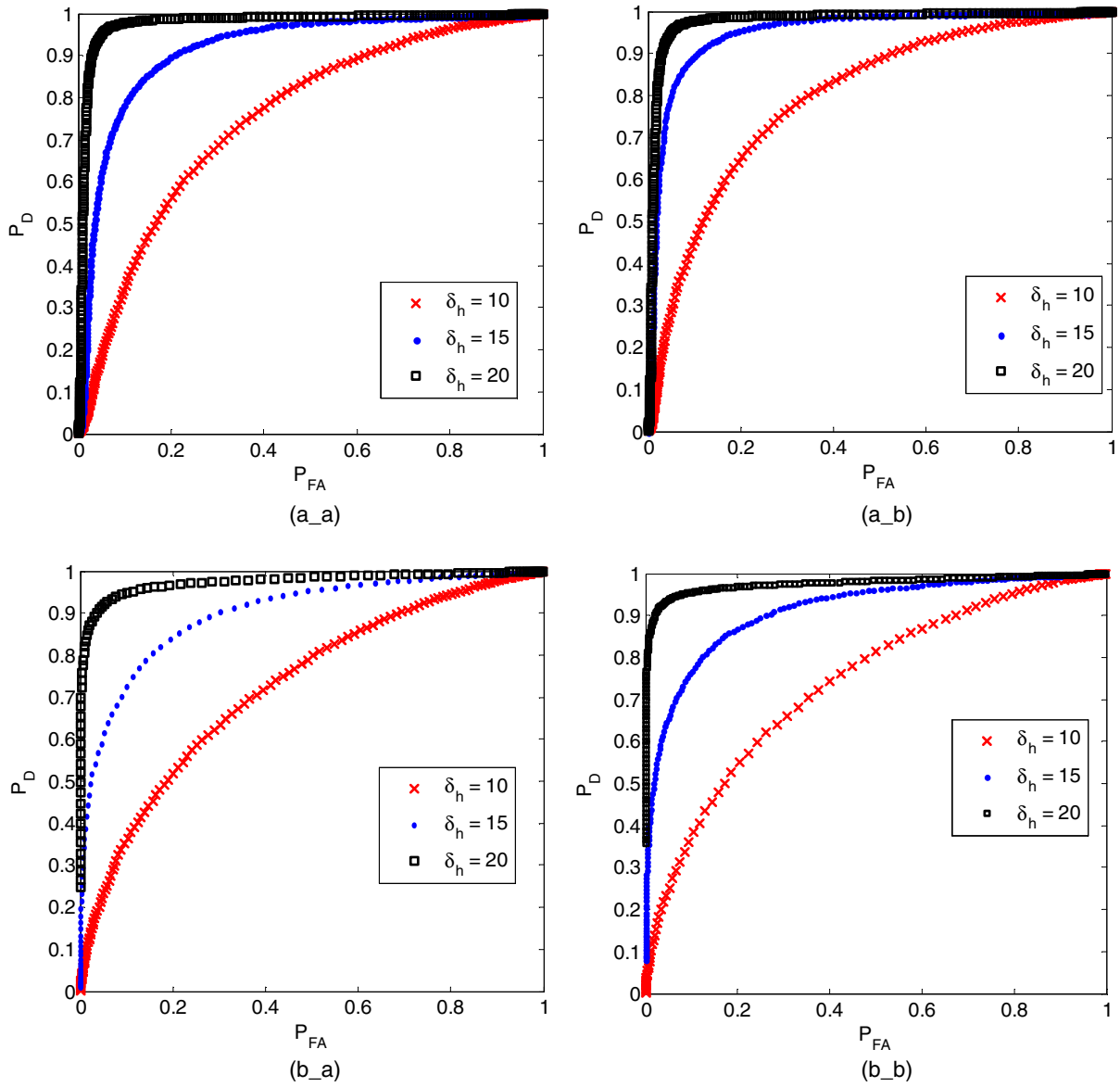

FIG. 11. (a) ROCs of locally optimum generalized likelihood ratio detector, $\mathrm{GLR}^{(\mathrm{LO})}$, Eq. (49), for different values of δ_h. Non Gaussian glitchy noise with $\mathrm{SNR}^{(g)}_{\max} = 10^2$ and $\lambda = 0.5\ \mathrm{s}^{-1}$. Top: linearly (+) polarized GWB, source at $\vartheta = 2.28$ rad, $\varphi = 1.99$ rad; bottom: linearly (×)-polarized GWB, source at $\vartheta = 2.28$ rad, $\varphi = 0.94$ rad. (b) ROCs of generalized cross-correlation detector, GCC, Eq. (53), for different values of δ_h. Non Gaussian glitchy noise with $\mathrm{SNR}^{(g)}_{\max} = 10^2$ and $\lambda = 0.5\ \mathrm{s}^{-1}$. Top: linearly (+) polarized GWB, source at $\vartheta = 2.28$ rad, $\varphi = 1.99$ rad; bottom: linearly (×)-polarized GWB, source at $\vartheta = 2.28$ rad, $\varphi = 0.94$ rad.

The detectors discussed in Secs. III and IV are designed for i.i.d. noises, and hence, strictly speaking, they are *sub* optimal in our m-dependent simulated noise. Our displayed results will thus give rather *conservative* estimates of the performance improvement obtainable using LO detectors in place of linear correlator based ones.

Before presenting them, we shall briefly discuss here some possible strategies to design nearly LO detectors in m-dependent (or more generally *colored*) non-Gaussian noises.

Locally optimum detection of deterministic signals in non-Gaussian m-dependent noise was perhaps first discussed by Poor and coworkers. In [148] it was shown that, in the usual threshold approximation, the structure of the LO detector is the same as for the i.i.d. case, except that now the optimal memoryless nonlinearity g_{LO} must be determined from the 2nd order distribution of the noise process, using suitable approximations (see, e.g., [149]).

In [150] a *lowest order* study of the effects of m dependence was exploited, using a moving-average model of m-dependent noise, expanding the detector performance measures in powers of the moving average coefficients, and retaining only the lowest (first) order terms. Under such approximation (*weak m* dependence), it was proved that the LO memoryless nonlinearity differs from the i.i.d. one by a

linear term, whose slope depends from the moving-average coefficients and the Fisher entropy of the generating i.i.d. process. For the case when this latter is only loosely known, robust criteria for determining the linear correction exist, which basically amounts to using the distribution rated as least-favorable in the i.i.d. case [150].

In a future paper, illustrating results based on real data, we shall implement our LO detectors using the spectral samples obtained from the data by discrete Fourier transform. Indeed, as stressed in [55], in the limit where the observational window is substantially longer than the noise correlation time, the (complex) spectral samples $\tilde{x}_k$ obtained by a discrete Fourier transformation of a non-Gaussian stationary process become *uncorrelated*. Neglecting higher order correlations, they can also be assumed as being *independent*, so that their (joint) probability density can be conveniently factorized, allowing to write the LO statistic in a simple form.

The coefficients in the detection statistics (49), (53) are direction dependent. Thus, even in the null hypothesis (no signal), the distribution of these statistics is DOA *dependent*, and the detection threshold corresponding to a fixed false alarm probability is DOA dependent, in turn.

This is illustrated in Fig. 4, where the detection thresholds for the statistics (49) and (53) are displayed as functions of the source position on the celestial sphere, for a false-alarm probability $P_{FA} = 10^{-2}$. The threshold value for each DOA was obtained from a sample of 10^4 different noise realizations.

The maps in Figs. 4–7 are computed from a grid of DOAs [151,152] in the celestial sphere, uniformly sampled in φ and $\cos\vartheta$.

Maps of the detection probability P_D vs DOA for linearly polarized GWBs at a false-alarm level $P_{\mathrm{FA}} = 10^{-2}$ are shown in Figs. 5–7. Also in this case the P_D value for each DOA was obtained from a sample of 10^4 different noise realizations. In all cases the injected waveform was a Sine-Gaussian GWB, whose strength was gauged by the *intrinsic* quantity

$$\delta_h = \left(\frac{2h_{rss}^2}{N_1}\right)^{1/2}, \tag{68}$$

where h_{rss} is defined in Eq. (16) and N_1 is the one-sided noise power spectral density.

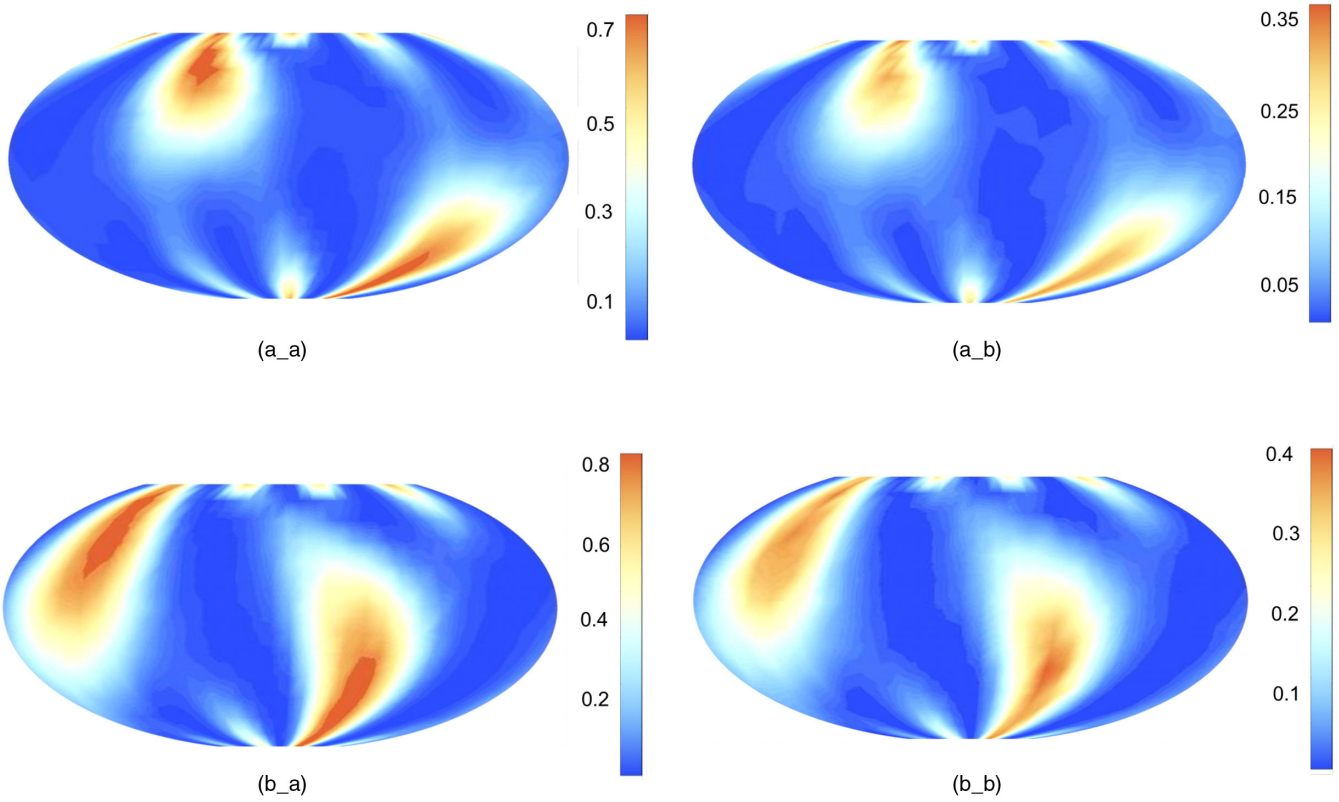

FIG. 12. (a) Mollweide maps of detection probability at a false alarm level of 10^{-2}. Data corrupted by non-Gaussian glitchy noise with $\mathrm{SNR}_{\mathrm{max}}^{(g)} = 10^2$ and $\lambda = 0.5\ \mathrm{s}^{-1}$. Linearly (+)-polarized GWB with $\delta_h = 20$. Top: robust version of GCC detector using the hard-limiter (HL) nonlinearity; bottom: nonparametric version of GCC detector using the signum nonlinearity. (b) Mollweide maps of detection probability at a false alarm level of 10^{-2}. Data corrupted by non-Gaussian glitchy noise with $\mathrm{SNR}_{\mathrm{max}}^{(g)} = 10^2$ and $\lambda = 0.5\ \mathrm{s}^{-1}$. Linearly (×)-polarized GWB with $\delta_h = 20$. Top: robust version of GCC detector using the hard-limiter (HL) nonlinearity; bottom: nonparametric version of GCC detector using the signum nonlinearity.

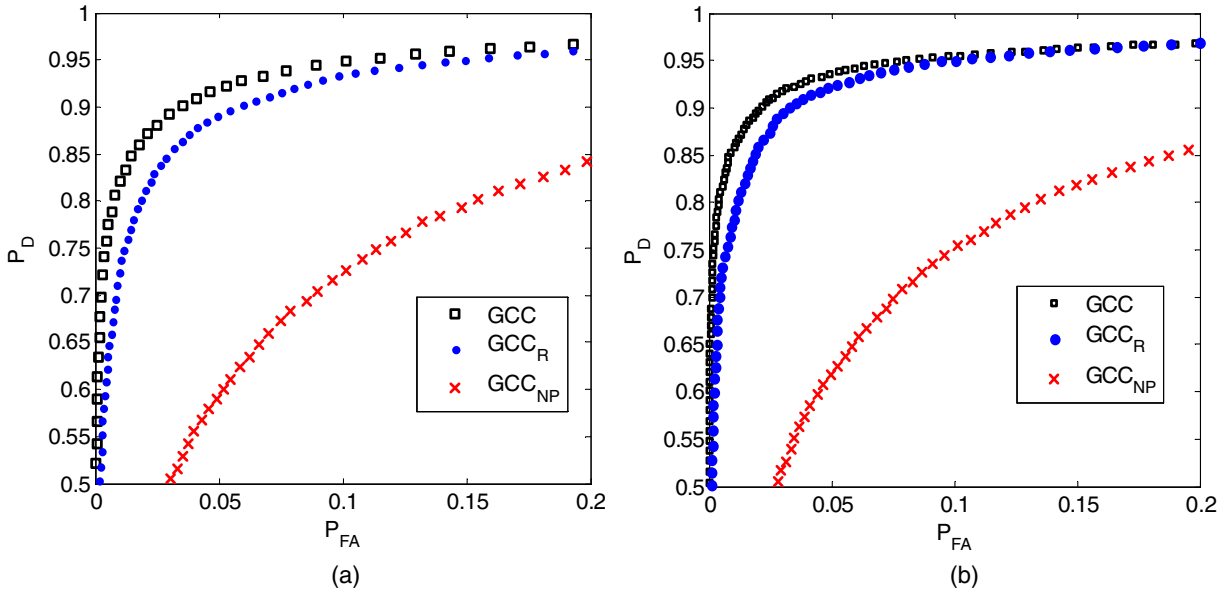

FIG. 13. (a) ROCs of generalized cross-correlation detector, Eq. (53), for different choices of the nonlinearity. Square markers: GCC detector ($g_{\rm LO}$ nonlinearity); dot markers: robust implementation GCC$_R$ (hard-limiter nonlinearity); cross markers: nonparametric implementation GCC$_{\rm NP}$ (signum nonlinearity). Non Gaussian noise with $\mathrm{SNR}^{(g)}_{\rm max} = 10^2$ and $\lambda = 0.5\ \mathrm{s}^{-1}$. Linearly (+)-polarized GWB with $\delta_h = 20$; source at $\vartheta = 2.28$ rad, $\varphi = 1.99$ rad. (b) ROCs of generalized cross-correlation detector, Eq. (53), for different choices of the nonlinearity. Square markers: GCC detector ($g_{\rm LO}$ nonlinearity); dot markers: robust implementation GCC$_R$ (hard-limiter nonlinearity); cross markers: nonparametric implementation GCC$_{\rm NP}$ (signum nonlinearity). Non Gaussian noise with $\mathrm{SNR}^{(g)}_{\rm max} = 10^2$ and $\lambda = 0.5\ \mathrm{s}^{-1}$. Linearly (×)-polarized GWB with $\delta_h = 20$; source at $\vartheta = 2.28$ rad, $\varphi = 0.94$ rad.

Note that the choice of the GWB *shape* is irrelevant (insofar as the waveform is entirely contained in the analysis window), in view of the very structure of the algorithms.

In Fig. 5, the data are corrupted by pure Gaussian noise, and the Gaussian-noise limiting forms of the (49) and (53) statistics are considered.

In Fig. 6, the same Gaussian-noise limiting form statistics are confronted with a glitchy-noise contaminated Gaussian background, with $\lambda = 0.5\ \mathrm{s}^{-1}$ and $\mathrm{SNR}^{(g)}_{\rm max} = 100$. A dramatic performance deterioration is observed, compared to Fig. 5. This is not unexpected (see, e.g., [153]).

In Fig. 7, the detection statistics (49) and (53) using the LO nonlinearity (23) are confronted with the same glitchy-noise contaminated Gaussian background, showing good performance recovery, by comparison with Fig. 6.

This is further illustrated in Fig. 8, in terms of receiver operating characteristics (ROCs, aka P_D vs $P_{\rm FA}$ curves, for fixed SNR). These curves refer to linearly polarized GWBs with $\delta_h = 20$, whose DOA correspond to the positions of maximum network sensitivity. The blue markers in Fig. 8 refer to the Gaussian-noise limiting forms of the detection statistics (49) and (53), and to data corrupted by pure Gaussian noise. The black markers refer to the same Gaussian-noise limiting forms of the detection statistics, confronted with a glitchy-noise contaminated Gaussian background, with $\lambda = 0.5\ \mathrm{s}^{-1}$ and $\mathrm{SNR}^{(g)}_{\rm max} = 100$. Finally, the red markers refer to the detection statistics (49) and (53) using the LO nonlinearity (23), in the same glitchy-noise contaminated Gaussian background.

The way ROCs are affected by varying the key factors affecting the glitchy noise component (namely, the glitch rate λ and the maximum signal to noise ratio $\mathrm{SNR}^{(g)}_{\rm max}$ of the glitches against the Gaussian floor), and/or the (intrinsic) signal to noise ratio δ_h is illustrated in Fig. 9, and in Figs. 10 and 11. In Figs. 10 and 11 only one GWB polarization is shown, for brevity.

It is seen that both detection statistics (49) and (53) using the locally optimum nonlinearity (23) outperform significantly their Gaussian limiting forms, when confronted with data corrupted by glitchy-noise contaminated Gaussian noise. Note that in the case where $\lambda = 1\ \mathrm{s}^{-1}$ and $\mathrm{SNR}^{(g)}_{\rm max} = 100$ and $\delta_h = 20$, the GW signal amplitude in LHO and LLO, falls beyond the knee point of the $g_{\rm LO}$ function, implying a trimming of the signal. Yet, the performance is still better than that of the linear correlator.

From the practical viewpoint, it is important to note that robust implementations of the same detectors, based on the HL nonlinearity discussed in Sec. V, does almost as well as those based on the *exact* LO nonlinearity. The signum nonlinearity, on the other hand, yields a poorer detector performance compared the LO and robust HL nonlinearities, but still outperforms the Gaussian-noise limiting form of these detectors.

This is illustrated in Figs. 12 and 13 (only the GCC is shown, for brevity).

VII. CONCLUSIONS AND HINTS FOR FUTURE WORK

Starting from the first principles, and resorting to the concept of local (weak signal) optimality, we discussed network detection of unmodeled GWB in non-Gaussian noise and validated the proposed detection strategies in simulated physically-inspired glitchy noise, whose parameters were obtained from the LIGO S5 data. Two possible network detectors of unmodeled GWB in non-Gaussian noise have been discussed, based respectively on the generalized likelihood ratio (GLR) and the generalized cross-correlation (GCC) statistics. Both detectors perform reasonably well in glitchy noise, without requiring *ad hoc* vetoing and/or data laundering preprocessing to remove high-SNR spurious glitches. The GLR and GCC statistics are straightforward generalizations of the correlators currently used in the LIGO-Virgo data analysis pipelines based on the assumption of Gaussian noise, the only difference being that each data sample should be filtered by a suitable (static) nonlinearity before computing the detection statistic. This latter retains the form of the usual linear correlator. The mentioned nonlinear transformations become linear when the noise is plain Gaussian, and the locally-optimum detectors merge into the familiar linear correlators.

The implementation of the proposed detectors would accordingly require *minimal* changes to the existing pipelines, and negligible added computational burden.

It is worth emphasizing that our approach takes into account, through the heavy tailedness of the noise distribution, the existence of loud as well as weak (individually undetectable) glitches, which *can* occur in conjunction with a GW signal, without posing a direct challenge of discriminating between instrumental and GW transients.

The GCC statistic, which treats the sought GWB as random processes, offers a good detection performance and easier implementation, requiring, at variance of the GLR, no (generalized) matrix inversion, and thus being exempt from ill-conditioning problems. On the other hand it provides no estimate of the detected waveform.

On a broader perspective, we also discussed, in an operational framework, the main non-Gaussian features of noise in gravitational wave interferometers, and suggested a physically driven statistical model for the impulsive (glitchy) component. Notably, only a few *gross* features (namely, the glitch rate and the glitch SNR distribution) of the glitchy component are relevant in shaping the first order noise probability density, whose knowledge is the only modeling information needed to compute the locally-optimum detectors. When the above noise features are incompletely specified, robust and/or nonparametric versions of the proposed detectors can be implemented, which use suitable approximate forms of the required nonlinearities. We discussed in some detail the Min-Max philosophy for designing locally-optimum detectors which are robust against glitch-rate nonstationarities. These detectors were compared in performance to those corresponding to a known noise PDF, and to their limiting form corresponding to Gaussian noise assumption.

Extensive simulation results have been presented. Remarkably, not only the ideal, exact locally optimum detectors, but also their robust (hard-limiter) and nonparametric (signum) implementations, outperform the Gaussian noise tailored linear correlator in the presence of glitches.

The analysis has been limited here to the triggered case (fiducially known source location). The more general case where the source location is unknown and should be estimated will be the subject of a forthcoming paper.

Tests based on real LIGO/Virgo data are being developed. Partial preliminary results [154] confirmed the performance improvement discussed here.

ACKNOWLEDGMENTS

Fruitful interaction with the LSC-burst group is gratefully acknowledged. Special thanks are due to B. Allen, L. Cadonati, J. Creighton, E. Katsavounidis, S. Klimenko, S. Mohanty, P. Saulson, P. Shawhan, and P. Sutton for stimulating discussions and encouragement at various stages of the development of this work. This work has been funded in part by the Italian Ministry for University and Scientific Research (MIUR Grant No. 20082J7FBN) and the Italian Institute for Nuclear Physics (INFN). Part of this work has been done by Maria Principe as a 2009/2010 Fulbright Visiting Student Researcher funded by the US-Italy Fulbright Commission, at the University of Texas (now University of Texas Rio Grande Valley) in Brownsville, TX 78520, USA.

APPENDIX: DERIVATION OF EQ. (52)

In this Appendix we include a derivation of the generalized cross-correlation statistic (53). We make the following working assumptions: (i) the noise samples, $\mathbf{n}_d(k) = n_d(t_k), \quad \forall\ k = 1, \ldots, N_s$, are i.i.d., with probability density function $f_n^{(d)}(x)$ having zero mean and variance σ_n^2; (ii) the GWB linear polarization components, $\mathbf{h}_+$ and $\mathbf{h}_\times$, are random independent process; (iii) the GWB and noise processes are independent. Accordingly, letting $\mathcal{V} = \{\mathbf{V}_1, \ldots, \mathbf{V}_D\}$ the output data matrix, we may write

$$\begin{cases} H_0 : f_{\mathcal{V}}(\mathcal{V}) = \prod_{k=1}^{N_s} \prod_{d=1}^{D} f_n^{(d)}(V_{dk}) \\ H_1 : f_{\mathcal{V}}(\mathcal{V}|\theta) = E\left[\prod_{k=1}^{N_s} \prod_{d=1}^{D} f_n^{(d)}(V_{dk} - \theta S_{dk})\right], \ \theta \neq 0 \end{cases}, \tag{A1}$$

where the expectation $E[\cdot]$ is taken with respect to the random GWB samples. For these latter we assume

$$E[S_{dk}]=0,\ \ \forall\, d=1,\ldots,D\ \ \forall\, k=1,\ldots,N_s$$
$$E[S_{dk}S_{pm}]=\mathcal{R}_{dp}\delta_{k-m},\ \ \forall\, d,p=1,\ldots,D\ \ \forall\, k,m=1,\ldots,N_s. \tag{A2}$$

Letting further

$$P(\theta)=\prod_{k=1}^{N_s}\prod_{d=1}^{D} f_n^{(d)}(V_{dk}-\theta S_{dk}), \tag{A3}$$

we may write:

$$f_{\mathcal{V}}(\mathcal{V})=P(0),\qquad f_{\mathcal{V}}(\mathcal{V}|\theta)=E[P(\theta)]. \tag{A4}$$

The LO detection statistic is obtained by differentiating $f_{\mathcal{V}}(\cdot|\theta)$ with respect to θ, and evaluating it at $\theta=0$. To first order,

$$\begin{aligned}\frac{d}{d\theta}f_{\mathcal{V}}(\mathcal{V}|\theta)&=E\left[\frac{dP(\theta)}{d\theta}\right]\\&=-E\left[\sum_{k=1}^{N_s}\sum_{d=1}^{D}P(\theta)\frac{\dot f_n^{(d)}(V_{dk}-\theta S_{dk})}{f_n^{(d)}(V_{dk}-\theta S_{dk})}S_{dk}\right],\end{aligned} \tag{A5}$$

which for $\theta=0$ becomes

$$\frac{d}{d\theta}f_{\mathcal{V}}(\mathcal{V}|\theta)\Big|_{\theta=0}=-P(0)\sum_{k=1}^{N_s}\sum_{d=1}^{D}E[S_{dk}]\frac{\dot f_n^{(d)}(V_{dk})}{f_n^{(d)}(V_{dk})}, \tag{A6}$$

which is identically zero in view of Eqs. (A2). The LO detection statistic is accordingly obtained from the second-order derivative of (A4) evaluated at $\theta=0$, viz.

$$\begin{aligned}\frac{d^2}{d\theta^2}f_{\mathcal{V}}(\mathcal{V}|\theta)\Big|_{\theta=0}&=P(0)\sum_{k=1}^{N_s}\sum_{d=1}^{D}E[S_{dk}^2]\frac{\ddot f_n^{(d)}(V_{dk})}{f_n^{(d)}(V_{dk})}\\&\quad+P(0)\sum_{k=1}^{N_s}\sum_{d=1}^{D}\sum_{m=1}^{N_s}\sum_{p=1}^{D}{}^{(p,m)\neq(d,k)}E[S_{dk}S_{pm}]\\&\quad\times\frac{\dot f_n^{(d)}(V_{dk})}{f_n^{(d}(V_{dk})}\frac{\dot f_n^{(p)}(V_{pm})}{f_n^{(p)}(V_{pm})}.\end{aligned} \tag{A7}$$

Hence in view of (A2),

$$\begin{aligned}\frac{d^2}{d\theta^2}f_{\mathcal{V}}(\mathcal{V}|\theta)\Big|_{\theta=0}&=P(0)\sum_{k=1}^{N_s}\sum_{d=1}^{D}E[S_{dk}^2]\frac{\ddot f_n^{(d)}(V_{dk})}{f_n^{(d)}(V_{dk})}\\&\quad+P(0)\sum_{k=1}^{N_s}\sum_{d=1}^{D}\sum_{\substack{p=1\\p\neq d}}^{D}\mathcal{R}_{dp}\\&\quad\times\frac{\dot f_n^{(d)}(V_{dk})}{f_n^{(d)}(V_{dk})}\frac{\dot f_n^{(p)}(V_{pk})}{f_n^{(p)}(V_{pk})},\end{aligned} \tag{A8}$$

which, apart for the irrelevant factor $P(0)$ reproduces the detection statistic (52).

[1] B. Abbott *et al.*, Observation of Gravitational Waves from Binary Black Hole Merger, Phys. Rev. Lett. **116**, 061102 (2016).

[2] B. Abbott *et al.*, GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence, Phys. Rev. Lett. **116**, 241103 (2016).

[3] J. Aasi *et al.*, Advanced LIGO, Classical Quantum Gravity **32**, 074001 (2015).

[4] D. V. Martynov *et al.*, Sensitivity of the advanced LIGO detectors at the beginning of gravitational wave astronomy, Phys. Rev. D **93**, 112004 (2016).

[5] J. G. Baker, M. Campanelli, F. Pretorius, and Y. Zlochower, Comparisons of binary black hole merger waveforms, Classical Quantum Gravity **24**, S25 (2007).

[6] F. Pretorius, Evolution of Binary Black Hole Spacetimes, Phys. Rev. Lett. **95**, 121101 (2005).

[7] M. Campanelli, C. O. Lousto, P. Marronetti, and Y. Zlochower, Accurate Evolutions of Orbiting Black-Hole Binaries Without Excision, Phys. Rev. Lett. **96**, 111101 (2006).

[8] U. Sperhake, E. Berti, and V. Cardoso, Numerical simulations of black-hole binaries and gravitational wave emission, C.R. Phys. **14**, 306 (2013).

[9] Harald P. Pfeiffer, Numerical simulations of compact object binaries, Classical Quantum Gravity **29**, 124004 (2012).

[10] N. Anderson *et al.*, The transient gravitational-wave sky, Classical Quantum Gravity **30**, 193002 (2013).

[11] B. Abbott *et al.*, Characterization of transient noise in advanced LIGO relevant to gravitational wave signal GW150914, Classical Quantum Gravity **33**, 134001 (2016).

[12] F. Acemese *et al.*, Advanced Virgo: a second-generation interferometric gravitational wave detector, Classical Quantum Gravity **32**, 024001 (2015).

[13] H. Luck *et al.*, The upgrade of GEO 600, J. Phys. Conf. Ser. **228**, 012012 (2010).

[14] Y. Aso, Y. Michimura, K. Somiya, M. Ando, O. Miyakawa, T. Sekiguchi, D. Tatsumi, and H. Yamamoto, Interferometer design of KAGRA gravitational wave detector, Phys. Rev. D **88**, 043007 (2013).

[15] C. S. Unnikrishnan, IndIGO and LIGO-INDIA: scope and plans for gravitational wave research and precision metrology in India, Int. J. Mod. Phys. D **22**, 1341010 (2013).
[16] P. Astone *et al.*, Methods and results of the IGEC search for burst gravitational waves in the years 1997–2000, Phys. Rev. D **68**, 022001 (2003).
[17] P. Astone *et al.*, Results of the IGEC-2 search for gravitational wave bursts during 2005, Phys. Rev. D **76**, 102001 (2007).
[18] L. Baggio, M. Cerdonio, I. S. Heng, A. Ortolan, G. A. Prodi, E. Rocco, G. Vedovato, and S. Vitale, IGEC toolbox for coincidence search, Classical Quantum Gravity **19**, 1541 (2002).
[19] N. Arnaud, M. Barsuglia, M.-A. Bizouard, P. Canitrot, F. Cavalier, M. Davier, P. Hello, and T. Pradier, Detection in coincidence of gravitational wave bursts with a network of interferometric detectors: geometric acceptance and timing, Phys. Rev. D **65**, 042004 (2002).
[20] A. Vicere, Optimal detection of burst events in gravitational wave interferometric observatories, Phys. Rev. D **66**, 062002 (2002).
[21] J. Sylvestre, Time-frequency detection algorithm for gravitational wave bursts, Phys. Rev. D **66**, 102004 (2002).
[22] N. Arnaud, M. Barsuglia, M.-A. Bizouard, V. Brisson, F. Cavalier, M. Davier, P. Hello, S. Kreckelbergh, and E. K. Porter, Coincidence and coherent data analysis methods for gravitational wave bursts in a network of interferometric detectors, Phys. Rev. D **68**, 102001 (2003).
[23] C. W. Helstrom, *Statistical Theory of Signal Detection*, (Pergamon Press, London, 1968).
[24] L. S. Finn, Aperture synthesis for gravitational-wave data analysis: deterministic sources, Phys. Rev. D **63**, 102001 (2001).
[25] Y. Gursel and M. Tinto, Near optimal solution to the inverse problem for gravitational-wave bursts, Phys. Rev. D **40**, 3884 (1989).
[26] L. Wen and B. F. Schutz, Coherent network detection of gravitational waves: the redundancy veto, Classical Quantum Gravity **22**, S1321 (2005).
[27] E. Flanagan and S. A. Hughes, Measuring gravitational waves from binary black hole coalescences. II. The waves' information and its extraction, with and without templates, Phys. Rev. D **57**, 4566 (1998).
[28] S. Klimenko, S. Mohanty, M. Rakhmanov, and G. Mitselmakher, Constraint likelihood analysis for a network of gravitational wave detectors, Phys. Rev. D **72**, 122002 (2005).
[29] M. Rakhmanov, Rank deficiency and tikhonov regularization in the inverse problem for gravitational-wave bursts, Classical Quantum Gravity **23**, S673 (2006).
[30] S. D. Mohanty, M. Rakhmanov, S. Klimenko, and G. Mitselmakher, Variability of signal to noise ratio and the network analysis of gravitational wave bursts, Classical Quantum Gravity **23**, 4799 (2006).
[31] T. Z. Summerscales, A. Burrows, L. S. Finn, and C. D. Ott, Maximum entropy for gravitational wave data analysis: inferring the physical parameters of core-collapse supernovae, Astrophys. J. **678**, 1142 (2008).
[32] I. Yakushin, Coherent all-sky search for gravitational wave bursts with LIGO, GEO and VIRGO detectors, LIGO Tech. Rep. LIGO-G070207-05-Z, 2007.
[33] K. Hayama, S. D. Mohanty, M. Rakhmanov, and S. Desai, Coeherent network analysis for triggered gravitational wave burst searches, Classical Quantum Gravity **24**, S681 (2007).
[34] S. Chatterji, A. Lazzarini, L. Stein, P. J. Sutton, A. Searle, and M. Tinto, Coherent network analysis technique for discriminating GW bursts from instrumental noise, Phys. Rev. D **74**, 082005 (2006).
[35] P. J. Sutton *et al.*, X-Pipeline: An analysis package for autonomous gravitational-wave burst searches, New J. Phys. **12**, 053034 (2010).
[36] A. Searle, P. J Sutton, M. Tinto, and G. Woan, Robust Bayesian detection of unmodelled bursts, Classical Quantum Gravity **25**, 114038 (2008).
[37] E. Thrane and M. Coughlin, Detecting Gravitational-Wave Transients at 5 Sigma: A Hierarchical Approach, Phys. Rev. Lett. **115**, 181102 (2015).
[38] J. B. Kanner, T. B. Littenberg, N. Cornish, M. Millhouse, E. Xhakaj, F. Salemi, M. Drago, G. Vedovato, and S. Klimenko, Leveraging waveform complexity for confident detection of gravitational waves, Phys. Rev. D **93**, 022002 (2016).
[39] T. B. Littenberg, J. B. Kanner, N. J. Cornish, and M. Millhouse, Enabling high confidence detections of gravitational-wave bursts, Phys. Rev. D **94**, 044050 (2016).
[40] R. Lynch, S. Vitale, R. Essick, E. Katsavounidis, and F. Robinet, An information theoretic approach to the gravitational-wave bursts detection problem, arXiv:1511.05955.
[41] S. Klimenko *et al.*, Method for detection and reconstruction of gravitational-wave transients with networks of advanced detectors, Phys. Rev. D **93**, 042004 (2016).
[42] A possible exception is represented by disturbances induced in widely distant interferometers by Schumann resonances see I. Kowalska *et al.*, Globally coherent short duration magnetic field transients and their effect on ground based gravitational-wave detectors, arXiv:1612.01102.
[43] N. J. Cornish and T. B. Littenberg, BayesWave: Bayesian inference for gravitational wave bursts and instrument glitches, Classical Quantum Gravity **32**, 135012 (2015).
[44] R. Penrose, A generalized inverse for matrices, Proc. Cambridge Philos. Soc. **51**, 406 (1955).
[45] L. K. Nuttall *et al.*, Improving the data quality of advanced LIGO based on early engineering run results, Classical Quantum Gravity **32**, 245005 (2015).
[46] B. Abbott *et al.*, Search for high frequency gravitational-wave bursts in the first calendar year of LIGO's fifth science run, Phys. Rev. D **80**, 102002 (2009).
[47] B. Abbott *et al.*, Search for gravitational-wave bursts in the first year of the fifth LIGO science run, Phys. Rev. D **80**, 102001 (2009).
[48] B. Abbott *et al.*, All-sky search for gravitational-wave bursts in the second joint LIGO-Virgo run, Phys. Rev. D **85**, 122007 (2012).
[49] T. B. Littenberg and N. J. Cornish, Separating gravitational wave signals from instrument artifacts, Phys. Rev. D **82**, 103007 (2010).

[50] T. B. Littenberg and N. J. Cornish, Bayesian inference for spectral estimation of gravitational wave detector noise, Phys. Rev. D **91**, 084034 (2015).
[51] T. T. Kadota, An optimum strategy for detection in the presence of random transient disturbance and white Gaussian noise, IEEE Trans. Inf. Theory **37**, 276 (1991).
[52] J. Powell, D. Trifirò, E. Cuoco, I. S. Heng, and M. Cavaglià, Classification methods for noise transients in advanced gravitational-wave detectors, Classical Quantum Gravity **32**, 215012 (2015).
[53] D. R. Cox and V. Isham, *Point Processes*, (CRC Press, Boca Raton, 2000).
[54] J. D. Creighton, Data analysis strategies for the detection of gravitational waves in non-Gaussian noise, Phys. Rev. D **60**, 021101 (1999).
[55] B. Allen, J. D. E. Creighton, É. É. Flanagan, and J. D. Romano, Robust statistics for deterministic and stochastic gravitational waves in non-Gaussian noise. I: frequentist analyses, Phys. Rev. D **65**, 122002 (2002).
[56] B. Allen, J. D. E. Creighton, É. É. Flanagan, and J. D. Romano, Robust statistics for deterministic and stochastic gravitational waves in non-Gaussian noise. II: Bayesian analyses, Phys. Rev. D **67**, 122002 (2003).
[57] M. Principe and I. M. Pinto, Modeling the impulsive noise component and its effect on the operation of a simple coherent network algorithm for unmodeled gravitational wave burst detection, Classical Quantum Gravity **25**, 075013 (2008).
[58] M. Principe and I. M. Pinto, Locally optimum network detection of unmodeled gravitational wave bursts in an impulsive noise background, Classical Quantum Gravity **26**, 045003 (2009).
[59] M. Principe and I. M. Pinto, Detecting unmodeled GW bursts in non-Gaussian (glitchy) noise: two locally optimum network detectors, Classical Quantum Gravity **26**, 204001 (2009).
[60] P. Rudnick, Likelihood detection of small signals in stationary noise, J. Appl. Phys. **32**, 140 (1961).
[61] M. Coughlin, Noise line identification in LIGO S6 and Virgo VSR2, J. Phys. Conf. Ser. **243**, 012010 (2010).
[62] P. Willems, V. Sannibale, J. Weel, and V. Mitrofanov, Mechanical Shot Noise Induced by Creep in Suspension Devices, Phys. Lett. A **297**, 37 (2002).
[63] The latter turns out to be negligible in fused silica suspensions [see I. A. Bilenko and N. Yu. Lyaskovskaya, The Investigation of Thermal and Non-Thermal Noises in Fused Silica Fibers for Advanced LIGO Suspensions, Phys. Lett. A **339**, 181 (2005)].
[64] A. M. Sintes and B. F. Schutz, Coherent line removal: filtering out harmonically related line interference from experimental data, with application to GW detectors, Phys. Rev. D **58**, 122003 (1998); Removing nonstationary non-harmonic external interference from gravitational wave interferometer data, Phys. Rev. D **60**, 062001 (1999).
[65] E. Chassande-Mottin and S. V. Dhurandhar, Adaptive filtering techniques for gravitational wave interferometric data: removing long-term sinusoidal disturbances and oscillatory transients, Phys. Rev. D **63**, 042004 (2001).
[66] E. J. Daw and M. R. Hewitson, A method for characterization of coherent backgrounds in real time and its applications in GW data analysis, Classical Quantum Gravity **25**, 205012 (2008).
[67] A. C. Searle, S. M. Scott, and D. E. McClelland, Spectral line removal in the LIGO data analysis system (LDAS), Classical Quantum Gravity **20**, S721 (2003).
[68] V. Tiwari *et al.*, Regression of environmental noise in LIGO data, Classical Quantum Gravity **32**, 165014 (2015).
[69] S. Mukherjee and L. S. Finn, Data conditioning for gravitational wave detectors: A Kalman filter for regressing suspension violin modes, Phys. Rev. D **63**, 062004 (2001); Erratum, Phys. Rev. D **67**, 109902(E) (2003).
[70] D. H. Santamore and Y. Levin, Eliminating thermal violin spikes from LIGO noise, Phys. Rev. D **64**, 042002 (2001).
[71] L. Blackburn *et al.*, The LSC glitch group: monitoring noise transients during the fifth LIGO science run, Classical Quantum Gravity **25**, 184004 (2008).
[72] J. Aasi *et al.*, Characterization of the LIGO detectors during their sixth science run, Classical Quantum Gravity **32**, 115012 (2015).
[73] S. Mukherjee (LIGO Collaboration), Preliminary results from the hierarchical glitch pipeline, Classical Quantum Gravity **24**, S701 (2007).
[74] S. Mukherjee, R. Obaid, and B. Matkarimov, Classification of glitch waveforms in gravitational wave detector characterization, J. Phys. Conf. Ser. **243**, 012006 (2010).
[75] S. Rampone, V. Pierro, L. Troiano, and I. M. Pinto, Neural network aided glitch-burst discrimination and glitch classification, Int. J. Mod. Phys. C **24**, 1350084 (2013).
[76] R. Biswas *et al.*, Application of machine learning algorithms to the study of noise artifacts in gravitational-wave data, Phys. Rev. D **88**, 062003 (2013).
[77] N. Mukund, S. Abraham, S. Kandhasamy, S. Mitra, and N. Philip, Transient classification in LIGO data using difference boosting neural network, arXiv:1609.07259.
[78] P. Ajith, M. Hewitson, J. R. Smith, and K. A. Strain, Robust vetoes for gravitational-wave burst triggers using known instrumental couplings, Classical Quantum Gravity **23**, 5825 (2006).
[79] P. Saulson, Listening to Glitches, LIGO Tech. Rep. LIGO-G070548-00-Z.
[80] J. R. Smith, T. Abbott, E. Hirose, N. Leroy, D. MacLeod, J. McIver, P. Saulson, and P. Shawhan, A hierarchical method for vetoing noise transients in gravitational-wave detectors, Classical Quantum Gravity **28**, 235005 (2011).
[81] P. Ajith, T. Isogai, N. Christensen, R. X. Adhikari, A. B. Pearlman, A. Wein, A. J. Weinstein, and B. Yuan, Instrumental vetoes for transient gravitational-wave triggers using noise-coupling models: The bilinear-coupling veto, Phys. Rev. D **89**, 122001 (2014).
[82] The philosophy of linear system theory, i.e., the description in terms of input-independent canonical responses and kernels, can be extended to (weakly) non linear systems using the Volterra-Wiener functional representation theory see, e.g. M. Schetzen, *The Volterra and Wiener Theory of Nonlinear Systems* (Wiley Interscience, New York, 2004).
[83] J. Powell, A. Torres-Forné, R. Lynch, D. Trifirò, E. Cuoco, M. Cavaglià, I.-S. Heng, and J. A. Font, Classification methods for noise transients in advanced gravitational-wave detectors II: performance tests on Advanced LIGO data, arXiv:1609.06262.

[84] J. Smith, Glitch classes seen in aLIGO so far (through end of O1), LIGO-Technical Report No. G1500642.
[85] D. Kontorovich and B. Lyandrev, Impulsive noise: a non-traditional approach, Signal Processing **51**, 121 (1996).
[86] M. L. Van Blaricum and R. Mittra, Problems and solutions associated with Prony's method for processing transient data, IEEE Trans. Antennas Propag. **26**, 174 (1978).
[87] V. Jain, T. Sarkar, and D. Weiner, Rational modeling by pencil-of-functions method, IEEE Trans. Acoust. Speech Signal Process. **6**, 897 (1981).
[88] P. Comon, Independent component analysis: a new concept?, Signal Processing **36**, 287 (1994).
[89] D. Middleton, On the theory of random noise. Phenomenological models. I, J. Appl. Phys. **22**, 1143 (1951); Erratum, J. Appl. Phys. **22**, 1326 (1951); On the theory of random noise. Phenomenological models. II, J. Appl. Phys. **22**, 1153 (1951).
[90] H. Hurwitz and M. Kac, Statistical analysis of certain types of random functions, Ann. Math. **15**, 173 (1944).
[91] S. D. Mohanty, Robust test for detecting nonstationarity in data from gravitational wave detectors, Phys. Rev. D **61**, 122002 (2000).
[92] M. Basseville and I. V. Nikiforov, *Detection of Abrupt Changes, Theory and Applications* (Prentice-Hall Inc., Englevood Cliffs, 1993).
[93] S. Mukherjee, Median-based noise floor tracker (MNFT): robust estimation of noise floor drifts in interferometric data, Classical Quantum Gravity **20**, S925 (2003).
[94] Partial evidence suggests (R. Conte, A SIRP model for LIGO Noise, PhD dissertation, University of Salerno (2008), Italy) that such a residual noise floor may be modeled as a compound-Gaussian spherically-invariant random process. See T. J. Barnard and D. D. Weiner, Non-Gaussian clutter modeling with generalized spherically invariant random vectors, IEEE Trans SP **44**, 2384 (1996), and references therein.
[95] H. Cramér, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, 1957).
[96] G. J. MacLachlan and D. Peel, *Finite Mixture Models* (J. Wiley & Sons, New York, 2000).
[97] This has been proved, in particular, for generalized shot noises belonging to the Middleton class see K. Vastola, Threshold detection in narrow-band non-Gaussian noise, IEEE Trans. Commun. **32**, 134 (1984); R. F. Ingram and R. Houle, Nav. Underw. Sys. Center (NUSC) Tech. Rep. 6339 (1980].
[98] R. DerSimonian, Algorithm AS 221: maximum likelihood estimation of a mixing distribution, J. R. Stat. Soc. **35**, 302 (1986).
[99] N. Vlassis and A. Likas, A greedy EM algorithm for Gaussian mixture learning, Neural Process. Lett. **15**, 77 (2002).
[100] S. Richardson and P. J. Green, On Bayesian analysis of mixtures with an unknown number of components (with discussion), J. R. Stat. Soc. Ser. B **59**, 731 (1997).
[101] R. S. Blum, Y. Zhang, and B. M. Sadler, On the approximation of correlated non-Gaussian noise PDFs using Gaussian mixture models, in *Proc. Conf. on Applications of Heavy Tailed Distributions in Economics, Engineering and Statistics* (American University, Washington DC 1999).
[102] D. Gabor, Theory of communication, J. Am. Inst. Electr. Eng. **93**, 429 (1946).
[103] M. J. Bastiaans, Gabor's expansion of a signal into elementary Gaussian signals, Proc. IEEE **68**, 538 (1980).
[104] S. Mallat and Z. Zhang, Matching pursuit with time-frequency dictionaries, IEEE Trans. Acoust. Speech Signal Process. **41**, 3397 (1993).
[105] The spectral width of SG atoms is set by their minimum-spread property $\sigma_f \sigma_t = (4\pi)^{-1}$.
[106] R. Viswanathan and P. K. Varshney, Distributed detection with multiple sensors: part I—fundamentals, Proc. IEEE **85**, 54 (1997).
[107] In principle, if the DOA is unknown, (ϑ_d, φ_s) can be regarded as unknown parameters, to be estimated according to the maximum likelihood principle, by picking up the largest statistic among those corresponding to a suitable lattice of points in the (ϑ_d, φ_s) space. Setting up in an optimal manner such a lattice requires knowledge of the related ambiguity function, which is left for a a future paper.
[108] The more realistic case of colored noise will be the subject of a future paper.
[109] M. Wąs, M.-A. Bizouard, V. Brisson, F. Cavalier, M. Davier, P. Hello, N. Leroy, F. Robinet, and M. Vavoulidis, On the background estimation by time slides in a network of gravitational wave detectors, Classical Quantum Gravity **27**, 015005 (2010).
[110] B. P. Abbott *et al.*, Observing gravitational wave transient GW150914 with minimal assumptions, Phys. Rev. D **93**, 122004 (2016).
[111] B. P. Abbott *et al.*, GW150914: First results from the search for binary black hole coalescence with advanced LIGO, Phys. Rev. D **93**, 122003 (2016).
[112] C. Messick *et al.*, Analysis framework for the prompt discovery of compact binary mergers in gravitational-wave data, Phys. Rev. D **95**, 042001 (2017).
[113] M. Wąs, M.-A. Bizouard, V. Brisson, F. Cavalier, M. Davier, P. Hello, N. Leroy, F. Robinet, and M. Vavoulidis, Limitations of the time slide method of background estimation, Classical Quantum Gravity **27**, 194014 (2010).
[114] B. Allen, χ^2 time-frequency discriminator for gravitational wave detection, Phys. Rev. D **71**, 062001 (2005).
[115] S. Babak *et al.*, Searching for gravitational waves from binary coalescence, Phys. Rev. D **87**, 024033 (2013).
[116] M. Wąs, P. J. Sutton, G. Jones, and I. Leonor, Performance of an externally triggered gravitational-wave burst search, Phys. Rev. D **86**, 022003 (2012).
[117] T. S. Adams, D. Meacher, J. Clark, P. J. Sutton, G. Jones, and A. Minot, Gravitational-wave Detection Using Multivariate Analysis, Phys. Rev. D **88**, 062006 (2013).
[118] T. Yamamoto, K. Hayama, S. Mano, Y. Itoh, and N. Kanda, Characterization of non-Gaussianity in gravitational wave detector noise, Phys. Rev. D **93**, 082005 (2016).
[119] P. Berry *et al.*, Parameter estimation for binary neutron-star coalescences with realistic noise during the advanced LIGO era, Astrophys. J. **804**, 114 (2015).
[120] S. A. Kassam, *Signal Detection in Non-Gaussian Noise* (Springer-Verlag, New York, 1988).
[121] E. H. Armstrong, A method of reducing disturbances in radio signaling by a system of frequency modulation, Proc. IRE **24**, 689 (1936).

[122] J. J. Lamb, A noise silencing I. F. circuit for superhet receivers, QST **2**, 11 (1936).
[123] J. H. Miller and J. N. Thomas, Robust detectors for signals in non-Gaussian noise, IEEE Trans. Commun. **25**, 686 (1977).
[124] J. B. Thomas, Nonparametric detection, Proc. IEEE **58**, 623 (1970).
[125] I. S. Heng, Rotating stellar core-collapse waveform decomposition: a principal component analysis approach, Classical Quantum Gravity **26**, 105005 (2009).
[126] A lower bound for P for GWBs of a given class could be obtained from the related information (Shannon) dimension, by using a well known property of quasi-bandlimited functions [D. Slepian, On Bandwidth, Proc. IEEE **64**, 292 (1976)]. Given the waveforms $\{S_p\}$, this amounts to computing $\sup_p \epsilon_p(B), \epsilon_p(B) = ||S_k(t) - S_k^{(B)}(t)||$, where $S_p^{(B)}(t) = \mathcal{F}^{-1}[\Pi_B(f)\mathcal{F}S_k(t)]$ where $||\cdot||$ is the L^2 norm on the finite time-support T_S of the $\{S_p\}$ family, $\mathcal{F}$ is the Fourier transform operator, and $\Pi_B(f)$ is the unit spectral window with support $(-B, B)$. The function $\epsilon_p(B)$ typically exhibits a steplike behavior, with knee point at $B = B_p^*$, and decays *exponentially* in $B - B_p^*$ afterwards. The information dimension of S_p can be accordingly estimated by $N_p \sim 2B_p^* T_S$, and $N_S = \sup_p N_p$ gives the estimate of the information dimension of the whole family of functions $\{S_p\}$.
[127] R. Maronna, R. D. Martin, and V. J. Yohai, *Robust Statistics: Theory and Methods* (Wiley, New York, 2006).
[128] A. M. Urmanov, A. V. Gribok, H. Bozdogan, J. W. Hines, and R. E. Uhrig, Information complexity-based regularization parameter selection for solution of ill-conditioned inverse problems, Inverse Probl. **18**, L1 (2002).
[129] P. J. Huber, Robust estimation of a location parameter, Ann. Math. **35**, 73 (1964).
[130] C. C. Lee and L. A. Longley, Nonparametric estimation algorithms based on input quantization, IEEE Trans. Inf. Theory **31**, 682 (1985).
[131] A. M. Maras, Threshold parameter estimation in non-additive non-Gaussian noise, IEEE Trans. Acoust. Speech Signal Process. **45**, 682 (1997).
[132] S. Marano, V. Matta, and P. Willett, Distributed estimation in large wireless sensor networks via a locally optimum approach, IEEE Trans. Acoust. Speech Signal Process. **56**, 748 (2008).
[133] M. Bertero, C. De Mol, and E. R. Pike, Linear inverse problems with discrete data: II. stability and regularisation, Inverse Probl. **4**, 573 (1988).
[134] M. Bertero, C. De Mol, and E. R. Pike, Linear inverse problems with discrete data: I—general formulation and singular system analysis, Inverse Probl. **1**, 301 (1985).
[135] As noted in [29], ill-conditioning will also amplify any error in the data due to inaccurate knowledge of the direction of arrival.
[136] The eigenvalues are real and positive, since the matrix $\mathbf{\Xi}^T\mathbf{\Sigma}\mathbf{\Xi}$ is symmetric and positive definite.
[137] S. A. Kassam, Locally Robust Array Detectors for Random Signals, IEEE Trans. Inf. Theory **24**, 309 (1978).
[138] R. D. Martin and S. C. Schwartz, Robust detection of a known signal in nearly Gaussian noise, IEEE Trans. Inf. Theory **17**, 50 (1971).
[139] Using (62) in (37) yields an estimator which minimizes the Huber norm of the residual of the linear system (29).
[140] A different robust detector was proposed in [J. W. Modestino, Adaptive detection of signals in impulsive noise environments, IEEE Trans. Commun. **25**, 1022 (1977)], assuming a *nominal* Gaussian stationary background corrupted by a *non-stationary* impulsive spurious signal. The statistic as a linear mixture of a signum-nonlinearity based threshold-type statistic, and of a linear (Gaussian noise tailored) correlator one, where the mixture coefficient is determined by maximizing the incremental SNR.
[141] More sophisticated nonparametric detectors are based on rank statistics, see, e.g., B. Seyfe and A. R. Sharafat, Nonparametric multiuser detection in non-Gaussian channels, IEEE Trans. Acoust. Speech Signal Process. **54**, 23 (2006); S. M. Zabin and G. A. Wright, Nonparametric density estimation and detection in impulsive interference channels. I. Estimators, IEEE Trans. Commun. **42**, 1684 (1994); Nonparametric density estimation and detection in impulsive interference channels. II. Detectors, IEEE Trans. Commun. **42**, 1698 (1994).
[142] The Laplace distribution represents a rather unphysical case, where the noise has infinite variance (and higher order moments). Heuristically, it is better regarded as a limiting case of (55), where $\epsilon \to 1$, and $K \to 0$. Accordingly, in order to implement the GLR, one may assume that $\Gamma(x)$ is a nonzero constant in some tiny neighbourhood of $x = 0$, whose actual value is irrelevant [it can be factored out in Eq. (35)].
[143] E. L. Lehman and J. P. Romano, *Testing Statistical Hypotheses* (Springer Verlag, New York, 2005).
[144] One should not be worried too much about the two limits in the ARE definition: they are needed in order to derive some formal properties of the ARE, but in practice the asymptotic value is already reached for $N_i \sim 100$ and SNR values of the order of 10, typically.
[145] J. Capon, On the asymptotic efficiency of locally optimum detectors, IRE Trans. Inf. Theory **7**, 67 (1961).
[146] R. E. Ziemer and R. B. Fluchel, Selection of blanking and limiting levels for binary signaling in gaussian plus impulse noise, in *Proc. of the 1971 IEEE Fall Electronics Conference (FEC '71)*, edited by G. R. Redinbo, Chicago, Oct. 18–20 1971, p. 290.
[147] This goes through the following steps [58]: (1) fix the glitch rate λ, select the (common) glitch waveform (atom), and assign the probability distributions of the glitch shape parameters (amplitude, center frequency, duration, etc.); (2) fix the width T of the analysis window, and draw the number of glitches in it, according to a Poisson law with parameter λ; (3) generate the random glitch firing times t_k, uniformly (and identically) distributed in T; (4) for each glitch draw the shape parameters values, using the chosen *a priori* distributions; (5) add white bandlimited Gaussian noise with specified variance.
[148] H. V. Poor and J. B. Thomas, Memoryless quantizer-detectors for constant signals in m-dependent noise, IEEE Trans. Inf. Theory **26**, 423 (1980).
[149] D. R. Halverson and G. L. Wise, A detection scheme for dependent noise processes, J. Franklin Inst. **309**, 287 (1980).

[150] H. V. Poor, Signal detection in the presence of weakly dependent noise-part I: optimum detection, IEEE Trans. Inf. Theory **28**, 735 (1982); Signal detection in the presence of weakly dependent noise–II: robust detection, IEEE Trans. Inf. Theory **28**, 744 (1982).

[151] The angular coordinates in Figs. 3–7 are the polar (ϑ) and azimuthal (φ) angles in an Earth-centered coordinate system, where the polar axis points to the North Pole and $\varphi = 0$ identifies the Prime Meridian.

[152] For DOAs near the poles, $\vartheta \to 0, \pi$, the threshold and the P_D displayed in Figs. 4–7 *vary* with φ. This is due to the fact that we keep the amplitude of the linearly polarized components of the incident wave *fixed*, i.e., DOA-independent. Note that the spherical unit vectors $(\hat{u}_\vartheta, \hat{u}_\varphi)$ are φ-dependent for $\vartheta \to 0, \pi$, so that our choice corresponds to a *differently oriented* source for each DOA. For a *fixed source*, the linearly polarized wave components as well as the antenna pattern functions would *oscillate* with φ in such a way that the instrument responses $\propto F_{+,\times} h_{+,\times}$ would be φ-independent in the limit $\vartheta \to 0, \pi$.

[153] J. C. Lindenlaub and K. A. Chen, Performance of matched filter receivers in non-Gaussian noise environments, IEEE Trans Commun Tech. **13**, 545 (1965).

[154] M. Principe, Noise modeling and reduction in GW detection experiments, PhD Dissertation, University of Sannio (2010).