Likelihood-ratio ranking of gravitational-wave candidates in a non-Gaussian background

Likelihood-ratio ranking of gravitational-wave candidates in a non-Gaussian background

Rahul Biswas,[1] Patrick R. Brady,[2] Jordi Burguet-Castell,[3] Kipp Cannon,[4] Jessica Clayton,[2] Alexander Dietz,[5] Nickolas Fotopoulos,[6] Lisa M. Goggin,[7] Drew Keppel,[8,9] Chris Pankow,[2] Larry R. Price,[6] and Ruslan Vaulin[10]

[1]*University of Texas-Brownsville, Brownsville, Texas 78520, USA*
[2]*University of Wisconsin–Milwaukee, Milwaukee, Wisconsin 53201, USA*
[3]*Universitat de les Illes Balears, E-07122 Palma de Mallorca, Spain*
[4]*Canadian Institute for Theoretical Astrophysics, University of Toronto, Toronto, Ontario, M5S 3H8, Canada*
[5]*The University of Mississippi, University, Mississippi 38677, USA*
[6]*LIGO - California Institute of Technology, Pasadena, California 91125, USA*
[7]*University of California San Francisco, San Francisco, California 94143 USA*
[8]*Albert-Einstein-Institut, Max-Planck-Institut für Gravitationsphysik, D-30167 Hannover, Germany*
[9]*Leibniz Universität Hannover, D-30167 Hannover, Germany*
[10]*LIGO - Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

(Received 20 February 2012; published 25 June 2012)

We describe a general approach to detection of transient gravitational-wave signals in the presence of non-Gaussian background noise. We prove that under quite general conditions, the ratio of the likelihood of observed data to contain a signal to the likelihood of it being a noise fluctuation provides optimal ranking for the candidate events found in an experiment. The likelihood-ratio ranking allows us to combine different kinds of data into a single analysis. We apply the general framework to the problem of unifying the results of independent experiments and the problem of accounting for non-Gaussian artifacts in the searches for gravitational waves from compact binary coalescence in LIGO data. We show analytically and confirm through simulations that in both cases applying the likelihood-ratio ranking results in an improved analysis.

DOI: 10.1103/PhysRevD.85.122008 PACS numbers: 04.80.Nn, 07.05.Kf, 95.55.Ym

I. INTRODUCTION

The detection of gravitational waves from astrophysical sources is a long-standing problem in physics. Over the past decade, the experimental emphasis has been on the construction and operation of kilometer-scale interferometric detectors such as Laser Interferometer Gravitational-wave Observatory (LIGO) [1]. The instruments measure the strain, $s(t)$, by monitoring light at the interferometer's output port, which varies as test masses that are suspended in vacuum at the ends of orthogonal arms differentially approach and recede by minuscule amounts. The strain signal, $s(t)$, is a combination of noise, $n(t)$, and gravitational-wave signal, $h(t)$.

There is a well established literature describing the analysis of time-series data for signals of various types [2]; these methods have been extended to address gravitational-wave detection [3]. This approach usually begins with the assumption that the detector noise, $n(t)$, is stationary and Gaussian. Then one proceeds to derive a set of filters that are tuned to detect the particular signals in this time-series data. The result is both elegant and powerful: whitened detector noise is correlated with a whitened version of the expected signal. The approach has been used to develop techniques to search for gravitational waves from compact binary coalescence, isolated neutron stars, stochastic sources, and generic bursts with certain time-frequency characteristics [4].

This approach takes the important first step of designing filters that properly suppress the dominant, frequency-dependent noise sources in the instrument. The simplicity of the filters is due to the fact that the power-spectral density fully characterizes the statistical properties of stationary, Gaussian noise. However, interferometric detectors are prone to non-Gaussian and nonstationary noise sources. Environmental disturbances, including seismic, acoustic, and electromagnetic effects, can lead to artifacts in the time series that are neither gravitational waves nor stationary, Gaussian noise. Imperfections in hardware can lead to unwanted signals in the time series that originate from auxiliary control systems.

To help identify and remove these unwanted signals, instruments have been constructed at geographically separated sites and the data are analyzed together. A plethora of diagnostics have also been developed to characterize the quality of the data [5–7]. Searches for gravitational waves use more than just the filtered output of the time-series, $s(t)$, to separate gravitational-wave signals from noise. Moreover, the responses from various filters indicate that the underlying noise sources are not Gaussian, even after substantial data quality filtering and coincidence requirements have been applied.

In this paper, we discuss using likelihood-ratio ranking as a unified approach to gravitational-wave data analysis. The approach foregoes the stationary, Gaussian model of the detector noise. The output of the filters derived under

that assumption becomes one element in a list of parameters that characterize a gravitational-wave detection candidate. The detection problem is then couched in terms of the statistical properties of an n-tuple of derived quantities, leading directly to a likelihood-ratio ranking for detection candidates. The n-tuple can include more information than simply the SNR measured in each instrument of the network. It can include measures of data quality, the physical parameters of the gravitational-wave candidate, the SNR from the coherent and null combinations of the detector signals; it can include nearly any measure of detector behavior or signal quality.

This approach was already used to develop ranking or detection statistics for compact binary coalescence signals [8–10] and is at the core of a powerful coincidence test developed for burst searches [11]. See also [12] for discussion of Monte-Carlo simulations and Bayesian techniques in searches for gravitational-wave bursts.

This work presents a general framework for the likelihood-ratio ranking in the context of gravitational-wave detection. We explore its analytical properties and illustrate its practical value by applying it to two data analysis problems arising in real-life searches for gravitational waves in LIGO data.

II. GENERAL DERIVATION OF LIKELIHOOD-RATIO RANKING

Let the n-tuple $\vec{c}$ denote the observable data related to one candidate event in some experiment that aims to detect a signal denoted by $\mathbf{h}$. This signal can usually be parametrized by several continuous parameters that may be unknown, for example, distance to the source of gravitational waves and its location on the sky. We impose no restriction on the nature of the observable data, $\vec{c}$. In particular, it can include readings from different detectors, information about the quality of data and the environment etc. The purpose of the experiment is to identify the signal. Depending on whether a Bayesian or frequentist statistical approach is taken, this is stated in terms of either the probability that the signal is present in the data or the probability that the observed data are a noise fluctuation.

In this section, we show that both approaches lead to ranking candidate signals according to the likelihood ratio

$$\Lambda(\vec{c}) = \frac{\int p(\vec{c}|\mathbf{h}, 1)p(\mathbf{h}|1)d\mathbf{h}}{p(\vec{c}|0)}, \tag{1}$$

where $p(\vec{c}|\mathbf{h}, 1)$ is the probability of observing $\vec{c}$ in the presence of the signal $\mathbf{h}$, $p(\mathbf{h}|1)$ is the prior probability to receive that signal, and $p(\vec{c}|0)$ is the probability of observing $\vec{c}$ in the absence of any signal. The higher a candidate's Λ value, the more likely it is a real signal.

We consider the most general case. As a consequence, we use an abstract notation for the key quantities, the observable data, $\vec{c}$, and a signal, $\mathbf{h}$, without specifying their precise nature. For clarity, it is worth to illustrate how they can be constructed using a real-life example. For this purpose we consider the problem of searching for a gravitational-wave signal from compact binary coalescence. The signal, $\mathbf{h}(\lambda_e, \lambda_i)$, in this case, is described by a set of intrinsic parameters, λ_i,–masses and spins of the compact objects, and a set of extrinsic parameters, λ_e–luminosity distance, location on the sky, orientation of the binary and time of coalescence. The starting point of the search for such a signal is to match-filter the time-series data, $s(t)$, with a bank of template waveforms which typically describe inspiral stage of coalescence of nonspinning compact objects and cover a range of masses expected to contain the targeted signal. The filters are derived using Gaussian model for the detector noise, $n(t)$, and include analytical maximization over the unknown extrinsic signal parameters, λ_e. The times when one of the template waveforms produce appreciable SNR, ρ, (usually exceeding a predefined threshold value) are recorded together with masses of the template waveform, (m_1, m_2), and constitute an initial list of the candidate events. Parameters of these candidate events, with exception of time which is noninformative, can be interpreted as reduced data, $\vec{c} \equiv (\rho, m_1, m_2)$. At the next stages, other quantities are computed for these candidate events, e.g. chi-square test for consistency with the signal waveform, difference in estimated time of arrival of the signal at different detectors etc., and the data vector, $\vec{c}$, can be extended. All of these parameters characterize candidate events and provide information that can be used to distinguish genuine gravitational-wave signals from non-Gaussian noise artifacts.

Some of the key questions one faces when trying to make use of this information are: What is the most optimal way of using the vector of parameters of the candidate event, $\vec{c}$, for signal detection? Can one construct the optimal ranking statistic for $\vec{c}$? Could detection efficiency of the search be reduced if one were to increase the mass range covered by the template waveforms or add another parameter characterizing candidate event and thus increase dimensionality of the detection problem? What if some parameters in $\vec{c}$ are useless for detection purpose, does using them imply less efficient search? If the noise or sensitivity of detector varies with time, what would be the optimal detection strategy in this case? Can one compare significance of the candidate events identified at different periods or experiments and combine their results?

In what follows we show that these questions can be made well defined and be answered within the likelihood-ratio framework. We should highlight, though, some of the subtle issues and often implicit assumptions that are commonly arise in practical applications and that are already contained in our example. First, note that the mass parameters, in our example, appear in both the signal, $\mathbf{h}$, and the observed data, $\vec{c}$. In the case of the signal, they are masses of compact objects in the binary, whereas in the case of the

observed data or the candidate event, they are masses of the template waveform that matched the data time series, $s(t)$, with high SNR. It is important to distinguish between the two sets of masses, because in general they do not coincide. The same logic applies to any other parameters that may appear in both $\mathbf{h}$ and $\vec{c}$. The space of signal parameters and the space of candidate event parameters are disjoined. Mapping between the spaces is induced by the conditional probability distribution, $p(\vec{c}|\mathbf{h}, 1)$.

Another important point to be aware of is that optimization is performed within the constraints set by the choice of the observed data, $\vec{c}$, and of the targeted signal, $\mathbf{h}$. The ideal data set would consist of all data points in the time series recorded by the detector and all other auxiliary data describing the state of the detector and the environment. In practice, one has to work with a reduced data set which inevitably results in some information loss. One of the tasks of a data analyst is to find the best possible reduced data set. One, then, can construct optimal ranking for the given reduced data that makes use of all available information.

Similarly, the choice of the targeted signals specified by $p(\mathbf{h}|1)$ constrains the process of optimization. The ideal choice for $p(\mathbf{h}|1)$ is the one that corresponds to the true astrophysical distribution of the parameters of the sources of gravitational waves (e.g. compact binary coalescence) in the universe, $p_{\text{astro}}(\mathbf{h}|1)$. Only in this case the ranking defined by the likelihood-ratio, Eq. (1), is optimal. In the absence of perfect knowledge of $p_{\text{astro}}(\mathbf{h}|1)$ one has to rely on an approximate model or a guess, $p_{\text{model}}(\mathbf{h}|1)$. Thus, strictly speaking, optimality is achieved only for the population of sources that match the model distribution, $p_{\text{model}}(\mathbf{h}|1)$, used to define the likelihood-ratio ranking. If the model distribution does not match the astrophysical one, the ranking is suboptimal. Nevertheless, unless the model distribution is completely wrong or internally inconsistent, this is the most reasonable starting point for data analysis. The model for $p_{\text{model}}(\mathbf{h}|1)$ can be corrected as more knowledge about astrophysical distribution is acquired. Everywhere in this paper, we assume that the true signal distribution is faithfully represented by the model distribution.

A. Bayesian analysis

In this approach, we compute the probability that a signal is present given the observed candidate event, $p(1|\vec{c})$. By a straightforward application of Bayes theorem, we write

$$p(1|\vec{c}) = \frac{p(\vec{c}|1)p(1)}{p(\vec{c}|1)p(1) + p(\vec{c}|0)p(0)} = \frac{\int p(\vec{c}|\mathbf{h}, 1)p(\mathbf{h}|1)p(1)d\mathbf{h}}{\int p(\vec{c}|\mathbf{h}, 1)p(\mathbf{h}|1)p(1)d\mathbf{h} + p(\vec{c}|0)p(0)}, \tag{2}$$

where $p(0)$ is the *prior* probability that the signal is absent and $p(1)$ is the *prior* probability that there is a signal (of any kind). These two outcomes are assumed to be mutually exclusive and exhaustive, $p(0) + p(1) = 1$. The denominator reexpresses $p(\vec{c})$ in terms of the two possible independent outcomes: the signal is present or the signal is absent. Upon successive division of numerator and denominator by $p(\vec{c}|0)$ and $p(1)$, we find

$$p(1|\vec{c}) = \frac{\Lambda(\vec{c})}{\Lambda(\vec{c}) + p(0)/p(1)}, \tag{3}$$

which is a monotonically increasing function of the likelihood ratio Λ defined by Eq. (1).[1] Hence, the larger the likelihood ratio, the more probable it is that a signal is present.

B. Frequentist approach

The process of detection can always be reduced to a binary ''yes'' or ''no'' question—does the observed data contain the signal? An optimal detection scheme should achieve the maximum rate of successful detections—correctly given yes answers—with some fixed, preferably low, rate of false alarms or false positives—incorrectly given yes answers. This is the essence of the Neyman-Pearson optimality criteria for detection, which states that an optimal detector should maximize the probability of detection at a fixed probability of false alarm [13].

As before, let the n-tuple $\vec{c}$ denote the observable data for a candidate event and $\mathbf{h}$ the signal that is the object of the search. Without loss of generality, any decision-making algorithm can be mapped into a real function, $f(\vec{c})$, of the data that signifies detection whenever its value is greater than or equal to a threshold value, F^*. Thus, using the Neyman-Pearson formalism, an optimal detector is realized by finding a function, $f(\vec{c})$, that maximizes the probability of detection at a fixed value of the probability of false alarm. The probability of detection, P_1, is

$$P_1 = \int_{V_d}\int_{V_{\mathbf{h}}} \Theta(f(\vec{c}) - F^*)p(\vec{c}|\mathbf{h}, 1)p(\mathbf{h}|1)p(1)d\mathbf{h}d\vec{c}, \tag{4}$$

and the probability of false alarm, [2] P_0, is

$$P_0 = \int_{V_d} \Theta(f(\vec{c}) - F^*)p(\vec{c}|0)p(0)d\vec{c}, \tag{5}$$

where $V_{\mathbf{h}}$ identifies the subset of signals targeted by the search, V_d denotes the subset of accessible data and integration is performed over all signals, $\mathbf{h}$, and data points, $\vec{c}$, within these subsets. Treating P_1 and P_0 as functionals of $f(\vec{c})$, we find that for an optimal detector, the variation of

$$S[f(\vec{c})] = P_1[f(\vec{c})] - l_0(P_0[f(\vec{c})] - P^*) \tag{6}$$

[1]This ratio of likelihoods is also known as the *Bayes factor.*

[2]This is similar, but not exactly equal, to the *false-alarm probability* or *Type I error*, which assumes the case where no signal is present, that is, does not include the term $p(0)$.

should vanish. Here l_0 denotes a Lagrange multiplier and P^* is a constant that sets the value of the probability of false alarm. The variation of Eq. (6) with respect to $f(\vec{c})$ gives

$$\delta S = \int_{V_d} \delta(f(\vec{c}) - F^*)\delta f(\vec{c})\Big[\int_{V_{\mathbf{h}}} p(\vec{c}|\mathbf{h}, 1)p(\mathbf{h}|1)p(1)d\mathbf{h} - l_0 p(\vec{c}|0)p(0)\Big]d\vec{c}. \quad (7)$$

Variations $\delta f(\vec{c})$ at different data points are independent, thus implying that after integration over $\vec{c}$, the condition

$$\frac{\int p(\vec{c}^*|\mathbf{h}, 1)p(\mathbf{h}|1)d\mathbf{h}}{p(\vec{c}^*|0)} = \frac{l_0 p(0)}{p(1)} = \text{const} \quad (8)$$

must be satisfied at all points $\vec{c}^*$ for which the argument of the delta function satisfies the condition

$$f(\vec{c}^*) - F^* = 0. \quad (9)$$

This latter condition defines the detection surface separating the detection and nondetection regions. Note that as $f(\vec{c})$ varies, the shape of this surface changes accordingly. Therefore, Eq. (8) implies that the optimal detection surface must be the surface of the constant likelihood ratio defined by the left hand side of the equation. This result is known as Neyman-Pearson Lemma [13] for simple hypothesis testing (e.g. signal with all its parameters known) and is generalizing it to the case of composite hypotheses (e.g. signal which parameters are not known or uncertain). It was also found earlier in [12], where a proof, albeit very different from ours, is outlined.[3] Eq. (8) is the only condition on the functional form of $f(\vec{c})$. Variation with respect to F^* does not give a new condition, whereas variation with respect to the Lagrange multiplier, l_0, simply sets the probability of a false alarm to be P^*.[4]

A natural way to satisfy the optimality criteria is to use the likelihood ratio

$$\Lambda(\vec{c}) = \frac{\int p(\vec{c}|\mathbf{h}, 1)p(\mathbf{h}|1)d\mathbf{h}}{p(\vec{c}|0)} \quad (10)$$

or any monotonic function $f(\Lambda(\vec{c}))$ for ranking the candidate signals. With this choice, the optimality condition Eq. (8) is satisfied for any threshold F^*. The latter is determined by the choice of an admissible value of the probability of false alarm, P^*, through

$$P_0[f(\Lambda(\vec{c}))] = P^*. \quad (11)$$

[3]We thank the referee for bringing to our attention this work that we have not been aware of at the time of working on the manuscript.

[4]In the case of the mixed data, when $\vec{c}$ includes continuous as well as discrete parameters, integration in the expressions for P_1 and P_0 should be replaced by summation wherever it is appropriate. This does not affect the derivation or the main result. The notion of optimal detection surface defined by Eq. (8) is straightforward to generalize to include both continuous and discrete data.

C. Variation of efficiency with volume of search space

The likelihood ratio defined by Eq. (1) is guaranteed to maximize the probability of signal detection for a given search. Because optimization is performed for a fixed region, V_d, in Eqs. (4) and (5), defined by all allowed values of a candidate's parameters, $\vec{c}$, in the search, it is unclear whether increasing the volume of available data (e.g. extension of the bank of template waveforms) would not result in an overall decrease of probability to detect signals. For example, one may be apprehensive of the potential increase in the rate of false alarms solely due to extension of the searched parameter space, V_d. Do not confuse this with possible expansion of the set of targeted signals, $V_{\mathbf{h}}$, on the space of signal parameters. Here we keep $V_{\mathbf{h}}$ fixed and vary only V_d. Intuitively, having more available information should not negatively affect the detection probability or efficiency if the information is processed correctly. In what follows, we prove that this is true if the likelihood ratio is used for making the detection decision.

To prove that the detection efficiency does not decrease when the range of candidate's parameters, $\vec{c}$, is increased, we must show that the variation of $\delta P_1/\delta V_d$ at a fixed P_0 is nonnegative. Consider a foliation of the space of data, V_d, by surfaces of constant likelihood ratio, S_Λ. Functionals for the probabilities of detection, Eq. (4), and of false alarm, Eq. (5) can be written as

$$P_1 = \int_0^\infty d\Lambda \int_{S_\Lambda} \Theta(\Lambda - \Lambda^*)p(\vec{c}|1)p(1)dS_\Lambda, \quad (12)$$

and

$$P_0 = \int_0^\infty d\Lambda \int_{S_\Lambda} \Theta(\Lambda - \Lambda^*)p(\vec{c}|0)p(0)dS_\Lambda, \quad (13)$$

where, for brevity, we absorbed explicit integration in the space of signals, $V_{\mathbf{h}}$, in the product $p(\vec{c}|1)p(1)$. P_1 is a functional of V_d and Λ^*. Since the latter is determined by the value chosen for false alarm probability, $P_0 = P^*$, and the probability of false alarm also depends on V_d, variations of V_d and Λ^* are not independent. To find the relation, we vary the probability of false alarm

$$\delta P_0 = -\int_0^\infty d\Lambda \int_{S_\Lambda} \delta(\Lambda - \Lambda^*)p(\vec{c}|0)p(0)\delta\Lambda^* dS_\Lambda + \int_0^\infty \Theta(\Lambda - \Lambda^*)p(\vec{c}|0)p(0)\delta S_\Lambda d\Lambda. \quad (14)$$

We consider nonnegative variations of surfaces of constant likelihood ratio, δS_Λ, that correspond only to the addition of new data points to V_d, and therefore correspond only to an extension of surfaces, S_Λ, without an overall translation or change of shape.

The probability of false alarm should stay constant, therefore its variation should vanish, providing the relation

$$\delta\Lambda^* = \frac{\int_0^\infty \Theta(\Lambda - \Lambda^*) p(\vec{c}|0)\delta S_\Lambda d\Lambda}{\int_{S_{\Lambda^*}} p(\vec{c}|0) dS_{\Lambda^*}}. \quad (15)$$

Next, we vary the functional for the detection probability

$$\delta P_1 = -\Lambda^* \int_{S_{\Lambda^*}} p(\vec{c}|0)p(1)\delta\Lambda^* dS_{\Lambda^*} + \int_0^\infty \Theta(\Lambda - \Lambda^*)\Lambda p(\vec{c}|0)p(1)\delta S_\Lambda d\Lambda, \quad (16)$$

where we use $p(\vec{c}|1) = \Lambda(\vec{c})p(\vec{c}|0)$, which follows from the definition of the likelihood ratio. Eliminating $\delta\Lambda^*$ by means of Eq. (15) and rearranging terms we get

$$\delta P_1 = p(1)\int_0^\infty \Theta(\Lambda - \Lambda^*)(\Lambda - \Lambda^*)p(\vec{c}|0)\delta S_\Lambda d\Lambda, \quad (17)$$

which is nonnegative for all positive δS_Λ by virtue of $\Theta(\Lambda - \Lambda^*)(\Lambda - \Lambda^*) \geq 0$. This proves that if the likelihood-ratio ranking is used in the detection process, the probability of detection can never decrease during an extension of the volume of available data.

D. Variation of efficiency with dimensionality of search space

Another way of changing the space of data, V_d, is by including a new parameter describing a candidate event to the n-tuple $\vec{c}$. This changes the dimensionality of V_d from n to $n + 1$. A new parameter may carry very important extra information about the candidate events which can help to distinguish true signals from noise artifacts, or it may be completely irrelevant in the context of detection. Whatever the case might be, one can show that incorporation of new data dimensions in the analysis, can never result in decrease of efficiency as long as the likelihood-ratio ranking is properly constructed and used. This statement agrees with the intuitive notion that having access to more data dimensions and, therefore, to more information should only improve the analysis.

Suppose the data tuple is extended by adding m new parameters, $\vec{c}' \equiv (\vec{c}, x_{n+1}, x_{n+2}, \ldots x_{n+m})$. The conditional probability distributions for observing $\vec{c}$ in presence or absence of a signal can be expressed in terms of corresponding probability distributions for $\vec{c}'$ marginalized over the added parameters:

$$p(\vec{c}|1) = \int p(\vec{c}', x_{n+1}, x_{n+2}, \ldots x_{n+m}|1) d^m\vec{x}, \quad (18a)$$

$$p(\vec{c}|0) = \int p(\vec{c}', x_{n+1}, x_{n+2}, \ldots x_{n+m}|0) d^m\vec{x}. \quad (18b)$$

Using Eqs. (18) the original probabilities of detection, Eq. (4), and of false alarm, Eq. (5) for detection problem in n-dimensional space of data can be written as integrals over extended $(n + m)$-dimensional data space:

$$P_1 = \int_{V_d'} \Theta(\Lambda(\vec{c}) - \Lambda^*) p(\vec{c}'|1)p(1)d\vec{c}', \quad (19)$$

and

$$P_0 = \int_{V_d'} \Theta(\Lambda(\vec{c}) - \Lambda^*) p(\vec{c}'|0)p(0)d\vec{c}', \quad (20)$$

where V_d' is extension of V_d by addition of the new parameters. Note that $\Lambda(\vec{c}) = p(\vec{c}|1)/p(\vec{c}|0)$ is a function of $\vec{c}$ only. Interpreting P_1 and P_0 as probabilities of detection and of false alarm for the search that uses $(n + m)$-dimensional data vector, $\vec{c}'$, and invoking the results of optimization analysis of Section II B, we conclude that $\Lambda(\vec{c})$ is either suboptimal or at best matches performance of the optimal ranking statistic in $(n + m)$-dimensional space, $\Lambda'(\vec{c}')$, in some special cases. This in turn implies that addition of new data dimensions can either improve the search, via applying more optimal ranking $\Lambda'(\vec{c}')$ in higher dimensional data space, or has no effect on efficiency. One of the special cases when addition of new parameters has no effect is situation when $\Lambda'(\vec{c}') = \Lambda(\vec{c})$. In other words, when the $(n + m)$-dimensional likelihood-ratio, Λ', does not vary with respect to any of the new parameters. It can happen when these parameters are either completely irrelevant to the detection problem or are functions of some (all) of the parameters in $\vec{c}$.

III. APPLICATIONS

In Sec. III A, we apply the formalism of Sec. II when assessing the significance of triggers between experiments on disjoint times. In Sec. III B, we demonstrate how the likelihood-ratio ranking can improve analysis efficiency by accounting for non-Gaussian features in the distributions of parameters of the candidate events.

A. Combining disjoint experiments

One complexity that arises in real-world applications is the necessity to combine results from multiple independent experiments. For example, gravitational-wave searches are often thought of in terms of times when a fixed number of interferometers are operating. If a network consists of instruments that are not identical and located at different places, each combination of operating interferometers may have very different combined sensitivity and background noise. Times when three interferometers are recording data may be treated differently from those when any pair is operating. Ideally, these experiments would be treated together accounting for differences in detectors' sensitivities and background noise in the ranking of the candidate signals, but it is often not practical (see how this problem was addressed in [10]). In this section, we show that the likelihood-ratio ranking offers a natural solution to this problem, which is conceptually similar to a simplified approach taken in [10].

Consider a situation in which the data is written as $\vec{c} = (\vec{d}, j)$, where $j = 0, 1, 2, \ldots$ indicates that the data arose from an experiment covering some time interval T_j. Note that $T_i \cap T_j = \varnothing$ if $j \neq i$. The probability that a signal is present given the data is

$$p(1|\vec{d}, j) = \frac{\int p(\vec{d}, j|\mathbf{h}, 1)p(\mathbf{h}|1)p(1)d\mathbf{h}}{\int p(\vec{d}, j|\mathbf{h}, 1)p(\mathbf{h}|1)p(1)d\mathbf{h} + p(\vec{d}, j|0)p(0)}. \quad (21)$$

The conditional probabilities for the observed data can be further expanded as

$$p(\vec{d}, j|\mathbf{h}, 1) = p(\vec{d}|j, \mathbf{h}, 1)p(j|\mathbf{h}, 1), \quad (22)$$

$$p(\vec{d}, j|0) = p(\vec{d}|j, 0)p(j|0), \quad (23)$$

where we introduce $p(j|\mathbf{h}, 1)$ and $p(j|0)$–the probabilities for a candidate event to belong to the jth experiment in presence or absence of a signal, respectively. It is reasonable to assume[5] that $p(j|\mathbf{h}, 1) = p(j|0)$, which implies that the time intervals for experiments, T_j, were defined without prior knowledge of when a signal is to occur. In this case, both probabilities drop out of Eq. (21), and the expression for the probability of a signal to be present in the data can be written as

$$p(1|\vec{d}, j) = \frac{\Lambda_j(\vec{d})}{\Lambda_j(\vec{d}) + p(0)/p(1)}, \quad (24)$$

with the likelihood ratio $\Lambda_j(\vec{d})$ given by

$$\Lambda_j(\vec{d}) = \frac{\int p(\vec{d}|j, \mathbf{h}, 1)p(\mathbf{h})d\mathbf{h}}{p(\vec{d}|j, 0)}. \quad (25)$$

Comparing Eq. (24) with Eq. (3), we conclude that the likelihood ratio $\Lambda_j(\vec{d})$, evaluated independently for each experiment, provides optimal unified ranking. In terms of their likelihood ratios, data samples from different experiments can be compared directly, with differences in experiments' sensitivities and noise levels being accounted for by $p(\vec{d}|j, \mathbf{h}, 1)$ and $p(\vec{d}|j, 0)$.

Following the steps outlined in Section II B, the same result can also be attained by direct optimization of the combined probability of detection at the fixed probability of false alarm. Optimality guarantees that the results of the less sensitive experiment can be combined with the results of the more sensitive experiment without loss of efficiency.

[5]This is not strict equality. Gravitational-wave events can alter the amount of live time in experiments to detect them. For example, an alert sounds in the LIGO and Virgo control rooms when gamma-ray bursts are detected, which sometimes accompany CBCs. The alert prompts operators to avoid routine maintenance and hardware injections, with their associated deadtimes, for the following 40 minutes.

In this approach, a unified scale provided by the likelihood ratio, $\Lambda_j(\vec{d})$, is explicit because, by construction, the same threshold, Λ^*, is applied to all data samples. Λ^* is determined by the value of the probability of false alarm for the combined experiment, given by

$$P_0 = \sum_j \int \Theta(\Lambda_j(\vec{d}) - \Lambda^*)p(\vec{d}|j, 0)p(j|0)p(0)d\vec{d}, \quad (26)$$

which makes the whole process less trivial. Notice that $p(j|0)$ (that can be approximated by $T_j/\sum_i T_i$) appears in the expression for P_0, however it does not appear in the expression for the likelihood ratio given in Eq. (25). Since $p(j|0)$ is proportional to the experiment duration, T_j, each experiment is weighted appropriately in the total probability of false alarm. In a similar fashion, experiment durations appear in the expression for the combined efficiency or the probability of detection for the combined experiment.

B. Combining search spaces

Sophisticated searches for gravitational-wave signals from compact binary coalescence [10,14–16] have been developed over the past decade. The non-Gaussian and nonstationary noise is substantially suppressed by the application of instrumental and environmental vetoes [5–7], coincidence between detectors, and numerous other checks on the quality of putative gravitational-wave signals. Nevertheless, the number of background triggers as a function of SNR depends on the masses of the binaries targeted in a search. For this reason, triggers have been divided into categories based on the chirp mass, $\mathcal{M}$, of the filter's template waveform that produced the trigger (where $\mathcal{M} = ((m_1 m_2)^3/(m_1 + m_2))^{1/5}$ and m_1 and m_2 are the masses of the compact objects in the binary). The background is a slowly varying function of $\mathcal{M}$, falling off more rapidly, as a function of SNR, for smaller values of $\mathcal{M}$. This is a manifestation of non-Gaussianities still present in the data. It is desirable to account for this dependence when ranking candidates found in the search.

In this section, we consider a toy problem that mimics the properties of the compact binary search but demonstrates how the likelihood-ratio ranking matches our intuition. Following that example, we present the results of a simulated compact binary search and demonstrate that the ranking statistic based on the likelihood ratio accounts for non-Gaussian features in background distribution and improves search efficiency.

1. Toy problem

Consider an experiment in which the data that define a candidate are $\vec{c} = (\rho, x)$, where ρ is the SNR and x is the extra parameter describing the data sample (e.g. the chirp mass of the template waveform). Suppose the distribution of the data in the absence of a signal is

$$p(\rho, x|0) = A\rho\exp(-\rho^2)\Theta(x)\Theta(1-x) + B\Theta(x+1)\Theta(-x)\Theta(\rho)\Theta(\alpha-\rho). \quad (27)$$

Figure 1 provides a graphic representation of this distribution. Notice that $p(\rho, x|0) = 0$ for $x < 0$ and $\rho > \alpha$, therefore data (ρ, x) in this region of the plane indicates the presence of a signal with unit probability. This intuition is clearly borne out in the above analysis since

$$p(\mathbf{h}|\rho, x) = \frac{p(\rho, x|\mathbf{h})p(\mathbf{h})}{p(\rho, x|\mathbf{h})p(\mathbf{h}) + 0} = 1 \quad (28)$$

for $\{(\rho, x)|x < 0 \text{ and } \rho > \alpha\}$, compare this equation with Eq. (2). The likelihood ratio for these data points is infinite, reflecting complete certainty that the data samples from this region are signals.

2. Simulated compact binary search

For the purpose of simulating a real-life search we use data from LIGO's fourth science run, February 24-March 24, 2005. The data was collected by three detectors: the H1 and H2 colocated detectors in Hanford, WA, and the L1 detector in Livingston, LA.

The search targets three types of binaries: neutron star-neutron star (BNS), neutron star-black hole (NSBH) and black hole-black hole (BBH). To model signals from these systems, we use nonspinning, post-Newtonian waveforms [17–27] that are Newtonian order in amplitude and second order in phase, calculated using the stationary phase approximation [18,25,26] with the upper cut-off frequency set by the Schwarzschild innermost stable circular orbit. We generate three sets of simulated signals, one for each type of binary. The neutron star masses are chosen randomly in the range $1–3M_\odot$, while the black hole masses are restricted such that the total binary mass is between $2–35M_\odot$. The maximum allowed distance for the source systems is set to 20 Mpc for BNS, 25 Mpc for NSBH and 60 Mpc for BBH. These distances roughly correspond to the sensitivity range of the detectors in this science run. All other parameters, including the location of the source on the sky, are randomly sampled. The simulated signals are

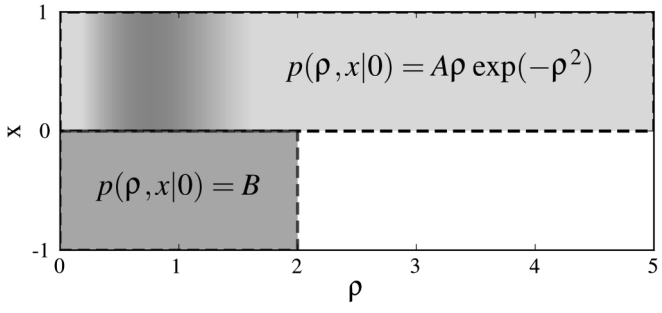

FIG. 1. Graphic representation of the model background distribution of Eq. (27) for $\alpha = 2.0$. Shaded areas define the regions of nonzero probability.

distributed uniformly in distance. In order to represent realistic astrophysical population with probability density function scaling as distance squared, the simulated signals are appropriately reweighted and are counted according to their weights. The simulated signals from each set are injected into nonoverlapping 2048-second blocks of data and analyzed independently.

Analysis of the data is performed using the low-mass CBC pipeline [10,14–16,28]. It consists of several stages. First, the time-series data recorded by each interferometer are match-filtered with the bank of nonspinning, post-Newtonian template waveforms covering all possible binary mass combinations with total mass in the range $2–35M_\odot$. The template waveforms come from the same family as the simulated signal waveforms previously described. When the SNR time series for a particular template crosses the threshold of 5.5, a single-interferometer trigger is recorded. This trigger is then subjected to waveform consistency tests, followed by consistency testing with triggers from the other interferometers. To be promoted to a gravitational-wave candidate, a signal is required to produce triggers with similar mass parameters in at least two interferometers within a very short time window (set by the light travel time between the detectors). The surviving coincident triggers are ranked according to the combined effective SNR statistic given by

$$\rho_c^2 = \sum_{i=1}^{N} \rho_{\mathrm{eff},i}^2, \quad (29)$$

where the sum is taken over the triggers from different detectors that were identified to be in coincidence and the phenomenologically constructed *effective SNR* for a trigger is defined as

$$\rho_{\mathrm{eff}}^2 = \frac{\rho^2}{\sqrt{(\frac{\chi^2}{2p-2})(1+\frac{\rho^2}{r})}}, \quad (30)$$

where ρ is the SNR, the phenomenological denominator factor $r = 250$, and p is the number of bins used in the χ^2 test, which is a measure of how much the signal in the data matches the template waveform [29]. In the denominator of Eq. (30), χ^2 is normalized by $2p - 2$, the number of degrees of freedom for this test.

All steps in the analysis beyond calculation of the SNR, ρ, are designed to remove non-Gaussian noise artifacts. Experience has shown that if properly tuned, these extra steps significantly reduce the number of false alarms [28]. Yet typically, the resulting output of the analysis is still not completely free of instrumental artifacts. Triggers that survived the pipeline's initial tests include unsuppressed noise artifacts. The general formalism developed in Sec. II can be applied to further classify these triggers with the aim of optimally separating signals from the noise artifacts. Each trigger is characterized by a vector of parameters which, in addition to the combined effective SNR, ρ_c,

may include the chirp mass, $\mathcal{M}$, difference in the time of arrivals at different detectors etc. Such information as which detectors detected the signal and what was the data quality at the time of detection can be also folded in as a discrete trigger parameter. For such parametrized data, the probability distributions in the presence and absence of a signal can be estimated via direct Monte-Carlo simulations. These distributions, if estimated correctly, include a non-Gaussian component. The triggers are ranked by their likelihood ratios, Eq. (1), which results in the optimized search in the parameter space of triggers.

Extra efficiency gained by additional processing of the triggers depends strongly on the extent to which the non-Gaussian features of the background noise are reflected in the distribution of the trigger parameters. In the context of the search for gravitational waves from compact binary coalescence in LIGO data, the chirp mass of a trigger's template waveform is one of the parameters that exhibits a nontrivial background distribution. For a given $\mathcal{M}$, the number of background triggers falls off with increasing combined effective SNR, ρ_c, of the trigger. The rate of falloff is slower for templates with higher chirp mass, reflecting the fact that non-Gaussian noise artifacts are more likely to generate a trigger for templates with smaller bandwidth. Another important piece of information about a trigger is the number and type of detectors that produced it. Generally, detectors differ by their sensitivities and level of noise. In the case we are concern with, two detectors, H1 and L1, have comparable sensitivities which are a factor of 2 higher than the sensitivity of the smaller H2 detector. This configuration implies that the signals within the sensitivity range of the H2 detector are likely to be detected in all three instruments forming a set of triple triggers, H1H2L1. The signals beyond the reach of the H2 detector can only be detected in two instruments forming a set of double triggers, H1L1. Detection of a true signal by another two detector combinations, H1H2 and H2L1, is very unlikely, therefore such triggers are discarded in the search. The number density of astrophysical sources grows as distance squared. As a consequence, it is more likely that a gravitational-wave signal is detected as an H1L1 double trigger. On the other hand, background of H1H2L1 triggers is much cleaner due to the fact that instrumental artifacts are less likely to occur in all three detectors simultaneously. These competing factors should be included in the ranking of the candidate events in order to optimize probability of detection.

It is natural to expect that inclusion of such information about the triggers in the ranking, in addition to the combined effective SNR, should help distinguishing signals from noise artifacts. The first step is to estimate distribution of trigger parameters for signals and background. For background estimation, we use the time-shifted data—the standard technique employed in the searches for transient gravitational-wave signals in LIGO data [10,14–16,28]. We perform 200 time shifts of the time-series data recorded by L1 with respect to the time-series data taken by the H1 and H2 detectors. The time lags are multiples of 5 seconds.

Analysis of time-shifted data provides us with a sample of the background distribution of the combined effective SNRs for H1L1 and H1H2L1 triggers with various chirp masses. We find that all triggers can be subdivided into three chirp mass bins: $0.87 \leq \mathcal{M}_c/M_\odot < 3.48$, $3.48 \leq \mathcal{M}_c/M_\odot < 7.4$, and $7.4 \leq \mathcal{M}_c/M_\odot < 15.24$. These correspond to equal mass binaries with total masses of 2–8$M_\odot$, 8–17$M_\odot$ and 17–35$M_\odot$. These same bins were used in the analyses of the data from LIGO's S5 and Virgo's VSR1 science runs [10,15,16]. Within each bin, the background distributions depend weakly on chirp mass, thus there is no need for finer resolution. At the same time, the distributions of the combined effective SNR in different bins show progressively longer tails with increasing chirp mass.

The distribution of triggers for gravitational-wave signals is simulated by injecting model waveforms into the data and analyzing them with the pipeline. This is done independently for each source type: BNS, NSBH and BBH.

Following the prescription for optimal ranking outlined in Sec. II, we treat each trigger as a vector of data $\vec{c} = (\rho_c, \alpha, m)$, where α denotes the type of the trigger, double H1L1 or triple H1H2L1, and m is a discrete index labeling the chirp mass bins. We construct the likelihood-ratio ranking, $\Lambda(\rho_c, \alpha, m|S_j)$ for each binary type, where S_j stands for BNS, NSBH or BBH. Note that the likelihood ratio has strong dependence on the binary type, S_j. To simplify calculations, we approximate the likelihood ratio by

$$\Lambda(\rho_c, \alpha, m|S_j) \approx \frac{n^j_{\rm inj}(\rho_c, \alpha, m)}{n_{\rm slide}(\rho_c, \alpha, m)}, \tag{31}$$

where $n^j_{\rm inj}(\rho_c, \alpha, m)$ is the fraction of injected signals of S_j type that produce a trigger of type α with $\rho'_c \geq \rho_c$ in the chirp mass bin m, and $n_{\rm slide}(\rho_{\rm eff}, \alpha, m)$ is the fraction of time shifts of the data that produce a trigger of type α with $\rho'_c \geq \rho_c$ in the same chirp mass bin. This approximation is equivalent to using cumulative probability distributions instead of probability densities. It is expected to be reasonably good for the tails of probability distributions that fall off as a power law or faster. The case we consider here falls into this category.

We compute the likelihood ratios given by Eq. (31) for all triggers: background and signals. Each trigger has three likelihood ratios, one for each binary type. We introduce a prior distribution for binary types, $p_s(S_j)$. It can either encode our knowledge about astrophysical populations of binaries or relative "importance" of different types of binaries to the search. In what follows we consider four alternatives: $p_s(S_j) = (1, 0, 0)$, $p_s(S_j) = (0, 1, 0)$, $p_s(S_j) = (0, 0, 1)$ and $p_s(S_j) = (1, 1, 1)$. The first three

singles out one of the binary types, whereas the last one treats all binaries on equal footing. Finally, the ranking statistic is defined as

$$\Lambda(\rho_c, \alpha, m) = \max_{S_j} \Lambda(\rho_c, \alpha, m|S_j) p_s(S_j). \quad (32)$$

Here we use maximization instead of marginalization over different types of signals, S_j, prescribed by the general form of the likelihood ratio, Eq. (1). This is a good approximation if signals of different types can be distinguished from each other by the analysis pipeline with high degree of confidence. Mathematically, it can be expressed as requirement that in the presence of a signal in the data $\Lambda(\rho_c, \alpha, m|S_j)$ is always strongly peaked on the correct type of signal. This is true in our case, because binary of different types produce triggers at different chirp-mass bins, m. Note though, that we do marginalize over signal parameters when computing likelihood ratio for each type of binary, S_j, in Eq. (31) by using population of simulated signals that samples the region in the parameter space pertained to the specific type of binary.

The four alternative choices for $p_s(S_j)$ define four different searches. For example, $p_s(S_j) = (1, 0, 0)$ corresponds to the search targeting only gravitational-wave signal from BNS coalescence. Similarly, $p_s(S_j) = (0, 1, 0)$ and $p_s(S_j) = (0, 0, 1)$ define the searches for gravitational waves from NSBH and BBH. The uniform prior, $p_s(S_j) = (1, 1, 1)$, allows one to detect all signals without giving priority to one type over the others. In each of the searches, the likelihood-ratio ranking, Eq. (32), reweights triggers giving higher priority to those that are likely to be the targeted signal as oppose to noise.

In order to assess the improvement attained by the new ranking, we compute efficiency in recovering simulated signals from the data as a function of the rate of false alarms. For a given rate of false alarms we find the corresponding value of the ranking and define efficiency as ratio of injected signals ranked above this value to the total number of signals that passed initial cuts of the analysis pipeline. This is equivalent to computing the standard receiver operating characteristic curve $P_1(P_0)$ defined by Eqs. (4) and (5). The efficiency curves are computed for BNS, NSBH and BBH binaries. In each case we evaluate efficiency of both likelihood-ratio rankings, the one that targets only that type of binary and the one that applies the uniform prior, $p_s(S_j) = (1, 1, 1)$. We compare the resulting curves to the efficiency curve for the standard analysis pipeline that uses the combined effective SNR, ρ_c as the ranking statistic. These curves are shown in Figs. 2 and 4.

They reveal that the searches targeting single type of binary, represented by the dashed curves, are more sensitive than the uniform search, the solid curve. This is expected, because narrowing down the space of signals typically allows one to discard the triggers that mismatch the signal's parameters reducing the rate of false alarms without loss of efficiency in recovering these signals. For instance, the search targeting BNS only signals discards all triggers from the high chirp-mass bins, $m = 2, 3$, without discarding the BNS signals. This reduces the rate of false alarms, although at the prices of missing possible gravitational-wave signals from other types of binaries, NSBH and BBH. Still, one could justify such search if it was known that NSBH and BBH binaries do not exist or are very rare. The uniform search, despite being less sensitive to BNS signals, allows one to detect the signals from all kinds of binaries. Such search still gains in efficiency over the standard search, the dotted curve, for BNS and NSBH systems, Figs. 2 and 3. At the same time, Fig. 4 shows that such search does worse in comparison to the standard search in detecting BBH signals. This is an unavoidable consequence of reweighting of triggers by the likelihood ratio, Eq. (32) based on their type and chirp mass. It ranks triggers from the lower chirp-mass bins higher, because these triggers are less likely to be a noise artifact. This leads to some loss of sensitivity to BBH signals, but gains sensitivity to BNS and NSBH signals. The role of the likelihood ratio is to provide optimal reweighting of triggers that results in the highest overall efficiency. In the case of the uniform search, it should provide increase in the total number of detected sources of all types. To demonstrate that this is indeed the case, we plotted the combined efficiency of the uniform search for

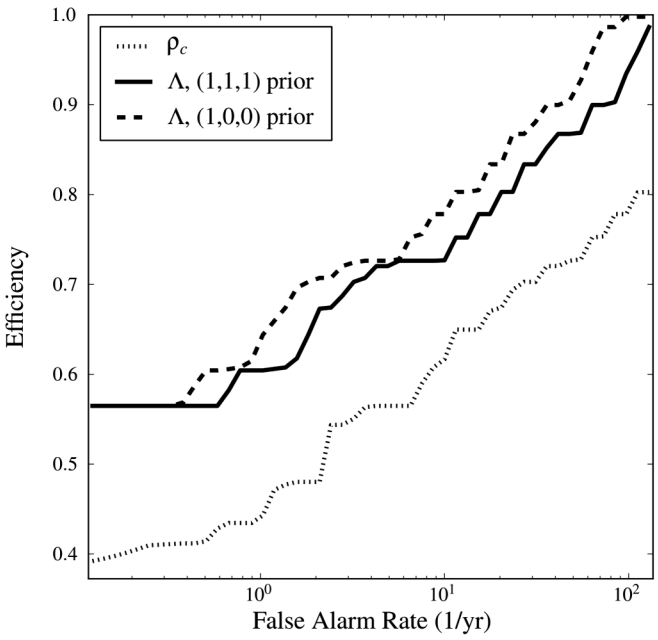

FIG. 2. Efficiency in detecting BNS signals versus false-alarm rate computed for various rankings. The solid curve corresponds to the likelihood-ratio ranking, Λ, with uniform prior $p_s(S_j) = (1, 1, 1)$. The dashed curve is the likelihood-ratio ranking, Λ, with the prior $p_s(S_j) = (1, 0, 0)$, singling out BNS binaries for detection. The dotted curve represents the standard search with the combined effective SNR ranking, $\rho_{\rm eff}$.

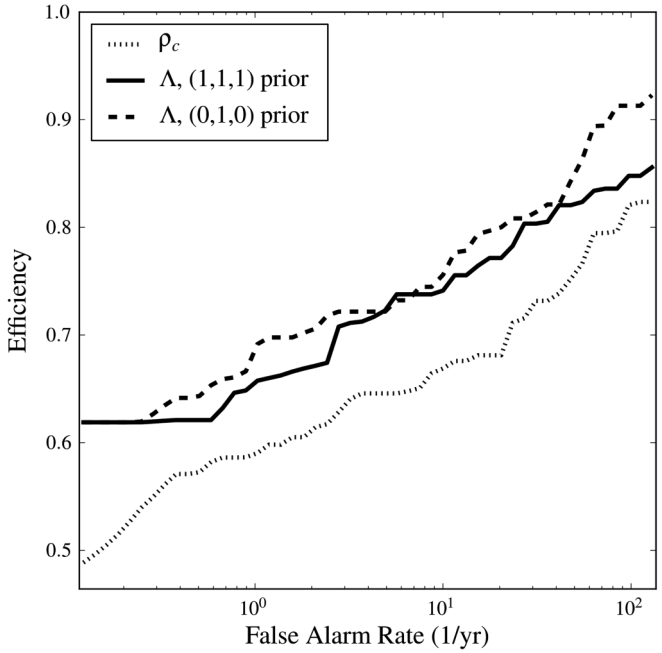

FIG. 3. Efficiency in detecting NSBH signals versus false alarm rate computed for various rankings. The solid curve corresponds to the likelihood-ratio ranking, Λ, with uniform prior $p_s(S_j) = (1, 1, 1)$. The dashed curve is the likelihood-ratio ranking, Λ, with the prior $p_s(S_j) = (0, 1, 0)$, singling out NSBH binaries for detection. The dotted curve represents the standard search with the combined effective SNR ranking, $\rho_{\rm eff}$.

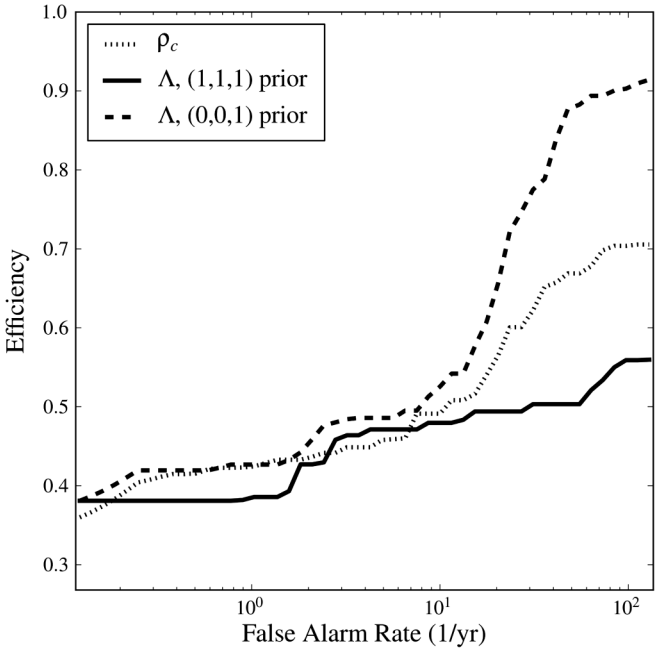

FIG. 4. Efficiency in detecting BBH signals versus false alarm rate computed for various rankings. The solid curve corresponds to the likelihood-ratio ranking, Λ, with uniform prior $p_s(S_j) = (1, 1, 1)$. The dashed curve is the likelihood-ratio ranking, Λ, with the prior $p_s(S_j) = (0, 0, 1)$, singling out BBH binaries for detection. The dotted curve represents the standard search with the combined effective SNR ranking, $\rho_{\rm eff}$.

BNS, NSBH and BBH signals and compared it to the efficiency of the standard search, Fig. 5.

The combined efficiency of the uniform search on Fig. 5 is higher than that of the standard search because triggers are reweighted by the likelihood ratio which properly accounts for the probability distributions of noise and signals. To gain further insight in this process we pick a particular point on the efficiency curve that corresponds to the rate of false alarms, x-axes, of 1.25 events per year. We find the corresponding to this rate threshold for combined effective SNR in the standard search to be $\rho_c^* = 11.34$. Next, we find the corresponding threshold for logarithm of the likelihood ratio, $\ln\Lambda^*(\rho_c, \alpha, m) = 9.11$. For each (α, m) combination this value can be mapped to ρ_c which will be different for each type of trigger. Both $\rho_c^* = 11.34$ and $\ln\Lambda^*(\rho_c, \alpha, m) = 9.11$ define detection surfaces in (ρ_c, α, m) space of trigger parameters. We depicted them on Fig. 6.

The signals falling to the right of $\rho_c^* = 11.34$, the dashed line, are considered to be detected in the standard search. Similarly, the signals that happen to produce a trigger to the right of $\ln\Lambda^*(\rho_c, \alpha, m) = 9.11$, the solid line, are considered to be detected in the uniform search with the likelihood-ratio ranking. The line of constant likelihood-ratio ranking sets different thresholds for combined effective SNR, ρ_c, of the triggers depending on their type. The threshold is higher than $\rho_c^* = 11.34$ for the H1L1 triggers

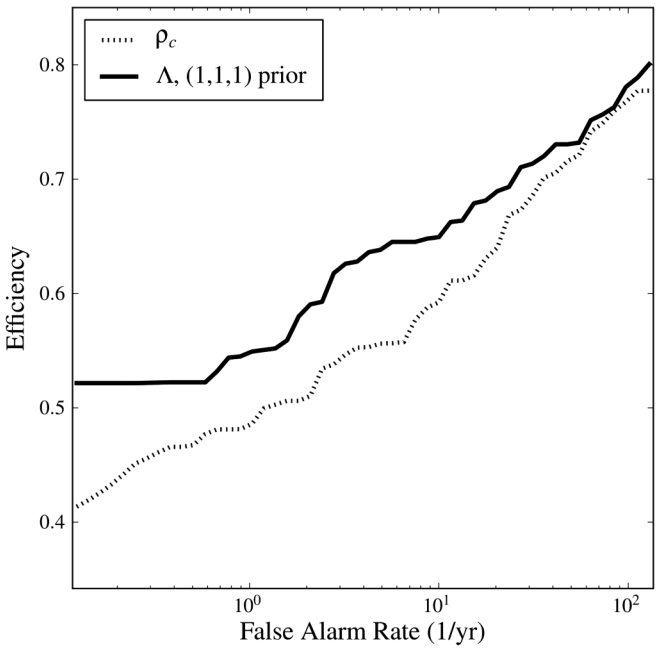

FIG. 5. Efficiency in detecting signals from any binary (BNS, NSBH or BBH) versus the false alarm rate computed for various rankings. The solid curve corresponds to the likelihood-ratio ranking, Λ, with uniform prior $p_s(S_j) = (1, 1, 1)$. The dotted curve represents the standard search with the combined effective SNR ranking, $\rho_{\rm eff}$.

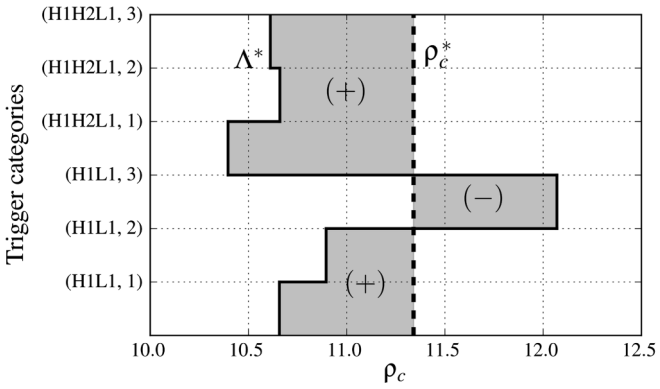

FIG. 6. The detection surfaces for the combined effective SNR, ρ_c, and the likelihood-ratio, Λ, rankings at the false alarm rate of 1.25 events per year. The y-axes labels different types and chirp-mass bins of the triggers. The dashed line is the line of constant combined effective SNR, $\rho_c^* = 11.34$. The solid line is the line of constant likelihood ratio, $\ln\Lambda^*(\rho_c, \alpha, m) = 9.11$. The signal producing a trigger that falls to the right of the dashed/solid curve is considered to be detected in the search with combined effective SNR/likelihood-ratio ranking. Those triggers that fall to the left are missed. The shaded region is the difference between the detection region for the likelihood ratio and the combined effective SNR rankings. The signals that produce trigger with parameters in the shaded regions labeled by $(+)$ are gained in the search equipped with likelihood ratio but missed by the search with the combined effective SNR ranking. Those signals that produce a trigger in the shaded region labeled by $(-)$ are missed by the likelihood-ratio ranking but detected by the combined effective SNR ranking.

from the third chirp-mass bin. The signals producing triggers in the shaded area in this bin, labeled by $(-)$, are missed in the uniform search but detected by the standard search. These signals are typically corresponds to BBH coalescence. The effect of this is visible on Fig. 4, the solid curve is below the dotted curve at false alarm rate of 1.25 events per year. On the other hand, the thresholds for other trigger types and chirp masses are lower than $\rho_c^* = 11.34$. As a result, the signals producing triggers with parameters in the shaded regions labeled by $(+)$ are detected in the uniform search but missed by the standard one. The net gain from detecting these signals is positive, Fig. 5. The process of optimization of the search in (ρ_c, α, m) parameter space can be thought of as deformation of the detection curve, $\rho_c^* = 11.34$, with the aim of maximizing efficiency of the search. The deformations are constrained to those that do not change the rate of false alarms. The optimal detection surface, as was shown in Sec. II B, Eq. (8), is the surface of constant likelihood ratio. This is the essence of likelihood ratio method.

The power of the likelihood-ratio ranking depends strongly on the input data. For demonstration purpose, in the simulation we restricted our attention to a subset of trigger parameters,(ρ_c, α, m). We expect that inclusion of other parameters such as difference in arrival times of the signal at different detectors, ratios of recovered amplitudes etc., should drastically improve the search. We leave this to future work.

IV. CONCLUSION

In this paper, we describe a general framework for designing optimal searches for transient gravitational-wave signals in data with non-Gaussian background noise. The principle quantity used in this method is the likelihood ratio, the ratio of the likelihood that the observed data contain signal to the likelihood that the data contain only noise. In Sec. II we prove that the likelihood ratio leads to the optimal analysis of data, incorporating all available information. It is robust against increase of the data volume, effectively ignoring irrelevant information. We apply the general formalism to two typical problems that arise in searches for gravitational-wave signals in LIGO data.

First, in Sec. III A we show that when searching for gravitational-wave signals in the data from different experiments or detector configurations, it is necessary as well as sufficient to rank candidates by the "local" likelihood ratio given by Eq. (25), which is calculated using estimated local probabilities. This provides overall optimality across the experiments. Candidate events from different experiments can be compared directly in terms of their likelihood ratios. This results in complete unification of the data analysis products. Another significant feature of the unified analysis is that the candidate's significance is independent of the duration of the experiments. Only the detectors' sensitivities and level of background noise contribute to the likelihood ratio of the candidates. The experiment's duration, on the other hand, measures its contribution to the total probability of detecting a signal (or efficiency) and the total probability of a false alarm.

Second, in Sec. III B we aim to improve efficiency of the search for gravitational waves from compact binary coalescence by considering the issue of consistent accounting for non-Gaussian features of the noise in the analysis. We suggest a practical solution to this problem. Estimate the probability distributions of parameters of the candidate events (e.g. SNR and the chirp mass of the template waveform, type of trigger etc.) in the presence and absence of a signal in the data. Construct the likelihood ratio that includes non-Gaussian features and use it to rerank candidate events. Nontrivial information contained in the probability distributions of candidate's parameters allows for a more optimal evaluation of their significance. Indeed, as we demonstrate in the simulation, inclusion of the chirp mass and the type of trigger in the likelihood-ratio ranking results in a significant increase of efficiency in detecting signals from coalescing binaries.

We would like to stress that the approach described in this paper is quite generic and can be applied to a wide range of problems in analysis of data with non-Gaussian background. Its main advantage is consistent account of

statistical information contained in the data. It provides a unified measure, in the form of the likelihood ratio, of the information relevant to detection of the signal in any type of data. This allows one to combine data of very different kind, such as the type of experiment, a type of trigger, its discrete and continuous parameters etc., into the single optimized analysis.

ACKNOWLEDGMENTS

Authors would like to thank Ilya Mandel and Jolien Creighton for many fruitful discussions and helpful suggestions. This work has been supported by NSF Grant Nos. PHY-0600953 and PHY-0923409. D. K. was supported from the Max Planck Gesellschaft. L. P. and R. V. was supported by LIGO laboratory. J. B. was supported by the Spanish MICINN FPA2010-16495 grant and the Conselleria dEconomia Hisenda i Innovacio of the Govern de les Illes Balears. LIGO was constructed by the California Institute of Technology and Massachusetts Institute of Technology with funding from the National Science Foundation and operates under cooperative agreement PHY-0757058.

[1] B. Abbott *et al.* (LIGO Scientific Collaboration), Rep. Prog. Phys. **72**, 076901 (2009).
[2] L. A. Wainstein and V. D. Zubakov, *Extraction of Signals from Noise* (Prentice-Hall, Englewood Cliffs, NJ, 1962).
[3] L. S. Finn and D. F. Chernoff, Phys. Rev. D **47**, 2198 (1993).
[4] W. G. Anderson, P. R. Brady, J. D. E. Creighton, and E. E. Flanagan, Phys. Rev. D **63**, 042003 (2001).
[5] N. Christensen (for the LIGO Scientific Collaboration and the Virgo Collaboration), Classical Quantum Gravity **27**, 194010 (2010).
[6] F. Robinet (for the LIGO Scientific Collaboration and the Virgo Collaboration), Classical Quantum Gravity **27**, 194012 (2010).
[7] D. M. Macleod, S. Fairhurst, B. Hughey, A. P. Lundgren, L. Pekowsky, J. Rollins, and J. R. Smith, Classical Quantum Gravity **29**, 055006 (2012).
[8] B. Abbott *et al.* (LIGO Scientific Collaboration), Astrophys. J. **681**, 1419 (2008).
[9] J. Abadie *et al.* (LIGO Scientific Collaboration), Astrophys. J. **715**, 1453 (2010).
[10] J. Abadie *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. D **82**, 102001 (2010).
[11] K. C. Cannon, Classical Quantum Gravity **25**, 105024 (2008).
[12] A. C. Searle, arXiv:0804.1161v1.
[13] J. Neyman and E. S. Pearson, Phil. Trans. R. Soc. A **231**, 289 (1933)
[14] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **77**, 062002 (2008).
[15] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **79**, 122001 (2009).
[16] B. Abbott *et al.* (LIGO Scientific Collaboration), Phys. Rev. D **80**, 047101 (2009).
[17] L. Blanchet, B. R. Iyer, C. M. Will, and A. G. Wiseman, Classical Quantum Gravity **13**, 575 (1996).
[18] S. Droz, D. J. Knapp, E. Poisson, and B. J. Owen, Phys. Rev. D **59**, 124016 (1999).
[19] L. Blanchet, Living Rev. Relativity **5**, 3 (2002).
[20] A. Buonanno, G. B. Cook, and F. Pretorius, Phys. Rev. D **75**, 124018 (2007).
[21] M. Boyle *et al.*, Phys. Rev. D **76**, 124038 (2007).
[22] M. Hannam *et al.*, Phys. Rev. D **77**, 044020 (2008).
[23] Y. Pan *et al.*, Phys. Rev. D **77**, 024014 (2008).
[24] M. Boyle, D. A. Brown, and L. Pekowsky, Classical Quantum Gravity **26**, 114006 (2009).
[25] K. S. Thorne, in *Three Hundred Years of Gravitation*, edited by S. W. Hawking and W. Israel (Cambridge University Press, Cambridge, 1987), Chap. 9, pp. 330–458
[26] B. S. Sathyaprakash and S. V. Dhurandhar, Phys. Rev. D **44**, 3819 (1991).
[27] B. J. Owen and B. S. Sathyaprakash, Phys. Rev. D **60**, 022002 (1999).
[28] B. Abbott *et al.* (LIGO Scientific Collaboration), Tech. Rep. No. LIGO-T070109-01, 2007.
[29] B. Allen, Phys. Rev. D **71**, 062001 (2005).