LIGO-P070076-00-Z

The Loudest Event Statistic: General Formulation, Properties and Applications

Rahul Biswas, Patrick R. Brady, and Jolien D. E. Creighton
University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA

Stephen Fairhurst
University of Wisconsin-Milwaukee, Milwaukee, WI 53201, USA
School of Physics and Astronomy, Cardiff University, Cardiff, CF2 3YB, United Kingdom and
LIGO - California Institute of Technology, Pasadena, CA 91125, USA

The use of the loudest observed event to generate statistical statements about rate and strength has been standard in searches for gravitational waves from compact binaries and pulsars. The Bayesian formulation of the method is generalized in this paper to allow for uncertainties both in the background estimate and in the properties of the population being constrained. The method is also extended to allow rate interval construction in the event of a very significant loudest event. Finally, it is shown how to combine the results from multiple experiments and a comparison is drawn between the upper limit obtained in a single search and the upper limit obtained by combining the results of two experiments each of half the original duration. To illustrate this, we look at an example case, motivated by the search for gravitational waves from binary inspiral.

I. INTRODUCTION

In daily life, we often estimate the birth rate, the rate of automobile fatalities, or the rate of hurricanes in the Gulf. In these cases, it is reasonably easy to determine when one event has occurred and so the best estimate is usually taken to be the number of events divided by the observation time. As physicists and astronomers, we know this is a good estimator of the rate of an underlying Poisson process. In these cases, the ability to identify events with high confidence is central to the correctness of the rate estimate.

We can carry this method over to more complicated observational situations by allowing for false positives in our identification of events. Experiments are usually designed so that the rate of real (foreground) events is higher than the rate of false positive (background) events. Hence a good estimate of the rate is obtained by counting the number of events per unit time, and making a small correction to allow for the false positives. This is the typical experimental method of estimating the rate.

In both physics and astronomy, it is common to search for very rare events in large data sets and we rely heavily on statistical methods to interpret these searches. In this paper, we discuss the problem of estimating the rate of these rare events. When real events are very rare or very weak, it is important to revisit the reasoning that underlies the standard approaches to estimating rates (and indeed other parameters). In this paper, we explore the effects of incorporating information about quality of observed events into the estimate of event rate. One measure of quality might be the signal to noise of the events; the louder an event, the more likely it is to be signal. Of course, more complicated measures are also possible. We simply require a rank ordering of the events such that larger quality implies less likely to be background.

A popular method of incorporating quality information is to fix a threshold, prior to looking at the data. The threshold is often chosen to give an acceptable rate of background events in some qualitative sense. Then, the upper limit is determined by counting the number of events per unit time above the chosen threshold and making a correction which allows for the background. Central to this method is the prescription by which the final list of events are identified.

There are many different criteria that might be used to determine the sample of events in an experiment. We consider using the loudest event to estimate the rate. This method was first introduced in gravitational-wave searches during the analysis data from a prototype instrument [1]; the method was used to determine an upper limit on the rate of binary neutron star mergers in the Galaxy. Since then, the method has been used in a number of searches for gravitational waves [2–7]. More details of this method of determining an upper limit are available in [8]. Related methods have been discussed in the context of particle physics experiments by Cousins [9] and Yellin [10].

In Sec. II, we present a general formulation of the loudest event statistic [1, 8]. We adopt the Bayesian approach which gives a posterior distribution over physical parameters based on the loudest event observed in an experiment. To provide a concrete example, in Sec. III we specialize to the case of a single unknown rate amplitude multiplying a known distribution of events. We discuss several properties of the statistic in more detail. In particular, we describe how the upper limit can be calculated in practice, appropriately taking into account the estimated background distribution. In later sections, we discuss marginalization over uncertainties, a method by which the loudest event statistic could be used in detections — by producing a confidence interval bounded away from zero, a method of combining the results for several independent experiments and the effect of splitting a single experiment in two halves. Throughout, we illustrate with an example where background events are Poisson distributed with a rate given by $\nu(x) = e^{-x^2/2}$ and the foreground is distributed at x^{-3}. These choices are natural in the context of a search for gravitational waves from coalescing binaries where the background is χ^2 distributed with two degrees of freedom, while the foreground is uniformly distributed in volume, and the signal strength (and hence statistic value x) are inversely proportional to the distance [11]. The main feature of these distributions, however, is simply that they are both decreasing functions of x, and that the background decreases more rapidly than the foreground.

II. GENERAL FORMULATION OF LOUDEST EVENT STATISTIC

Consider a search of experimental data for a rare process. The output of this search is a set of candidate events which have survived all cuts applied during the analysis. Assume that these events can be ranked according to a single parameter x where the chance of the event being due to instrumental noise decreases with increasing x. Suppose, further that we can parametrize the rate of events which occur and produce an event with a statistic value above x as $R(x)$. Then, if we assume that events are Poisson distributed in time, the probability of observing no events above a given value of x is

$$P(x) = e^{-R(x)} \,. \tag{1}$$

A real experiment will have a background of events due to noise. We assume that this can be determined. We will denote the probability of obtaining zero background events at or above a value x as $P_b(x)$. Then, taking into account the background, the probability of there being no events louder than x is

$$P(x|B) = P_b(x) e^{-R(x)} \tag{2}$$

where we have used B to indicate that the probability depends upon the background.

In an observation, the rate of events $R(x)$ will depend upon many quantities, for example the details of the apparatus used, the methods used in performing the search, a physical or astrophysical model for the source strength and distribution. We split these effects into two separate categories: those which depend upon the details of the search, and those which depend upon the physical or astrophysical process whose rate we are interested in measuring. The sensitivity of the search is encoded in the efficiency, ϵ, which is the probability of detecting a given source with a value greater than or equal to x. Clearly, it will depend upon the statistic value x, as well as a set of parameters, denoted $\mathcal{A}$, describing the amplitude of the signal emitted by the candidate source. Thus, we can write $\epsilon = \epsilon(x, \mathcal{A})$. For example, in the case of a search for gravitational waves from coalescing binaries, the efficiency will depend upon the sensitivity of the detector, the masses of the binary components, the sky location and orientation of the binary and its distance from the detector. These will all be encoded in the variable $\mathcal{A}$.

In addition, we require the physical rate $\mathcal{R}$ for sources with a given set of parameters $\mathcal{A}$. We assume that $\mathcal{R}$ can be parametrized by one or more physical parameters $\boldsymbol{\lambda}$ which we are interested in measuring or placing limits on. Thus, we can write the physical rate as $\mathcal{R}(\mathcal{A}|\boldsymbol{\lambda})$. Combining the above, it follows that

$$R(x|\boldsymbol{\lambda}) = \int d\mathcal{A} \; \epsilon(x, \mathcal{A}) \; \mathcal{R}(\mathcal{A}|\boldsymbol{\lambda}) \,. \tag{3}$$

Thus, we see that the rate of events in a search can be parametrized in terms of the model parameters $\boldsymbol{\lambda}$ and the threshold value x. Furthermore, it is broken into two distinct parts: the efficiency ϵ which depends upon the instrumental sensitivity and the details of the search; the physical rate $\mathcal{R}$ which depends upon the physical model used. Note that if the efficiency is independent of any of the physical parameters $\mathcal{A}_i$ then they can be trivially integrated out.

We can substitute our expression for the rate (3) into Eq. (2) to obtain the probability that there are zero events in the data with a statistic value greater than x as

$$P(x|\boldsymbol{\lambda}, \epsilon, B) = P_b(x) e^{-\int d\mathcal{A} \; \epsilon(x,\mathcal{A}) \; \mathcal{R}(\mathcal{A}|\boldsymbol{\lambda})} \,. \tag{4}$$

Furthermore, the probability of the loudest event occurring between x and $x + dx$ is given by $p(x|\boldsymbol{\lambda}, \epsilon, B)\, dx$ where

$$p(x|\boldsymbol{\lambda}, \epsilon, B) = P'(x|\boldsymbol{\lambda}, \epsilon, B) \tag{5}$$

and the prime denotes a derivative with respect to x.

After performing an experiment, we are interested in obtaining a distribution for the rate. To do this, we calculate a Bayesian posterior distribution for the parameters $\boldsymbol{\lambda}$ given the observations. This is denoted, $p(\boldsymbol{\lambda}|x_{\rm m}, \epsilon, B)$, where $x_{\rm m}$ is the observed loudest event. This distribution can be derived, using Bayes' law, from the prior probability distribution on the model parameters $p(\boldsymbol{\lambda})$ and the distribution in Eq. (5) as

$$p(\boldsymbol{\lambda}|x_{\rm m}, \epsilon, B) = \frac{p(\boldsymbol{\lambda}) \; p(x_{\rm m}|\boldsymbol{\lambda}, \epsilon, B)}{\int d\boldsymbol{\lambda} \; p(\boldsymbol{\lambda}) \; p(x_{\rm m}|\boldsymbol{\lambda}, \epsilon, B)} \,. \tag{6}$$

The denominator is required to ensure the probability distribution is normalized to unity.

More explicitly, we have

$$p(\boldsymbol{\lambda}|x_{\rm m}, \epsilon, B) \propto p(\boldsymbol{\lambda}) \left[\left(\frac{p_b(x_{\rm m})}{P_b(x_{\rm m})} \right) + \int d\mathcal{A} \; |\epsilon'(x_{\rm m}, \mathcal{A})| \; \mathcal{R}(\mathcal{A}|\boldsymbol{\lambda}) \right] e^{-R(x_{\rm m}|\boldsymbol{\lambda})} \,. \tag{7}$$

Here, $p_b(x) = P_b'(x)$ and, as before, the prime denotes derivative with respect to x. We have made use of the fact that the efficiency can only decrease as the value of x increases so that $\epsilon' \leq 0$. Looking at Eq. (7), notice that the probability distribution is governed by an exponential decay determined by the rate of events. The shape of the distribution is governed by two terms, the first $p_b(x_\mathrm{m})/P_b(x_\mathrm{m})$ depends only upon the background, while the second depends only upon the rate.

In many circumstances, the parameters $\boldsymbol{\lambda}$ may be further divided into a set of particular interest $\boldsymbol{\lambda}_0$ and others of less interest $\boldsymbol{\lambda}_1$. By integrating Eq. (6) over the unwanted parameters $\boldsymbol{\lambda}_1$, one obtains the posterior distribution

$$p(\boldsymbol{\lambda}_0|x_\mathrm{m},\epsilon,B) = \frac{\int d\boldsymbol{\lambda}_1\, p(\boldsymbol{\lambda}_0,\boldsymbol{\lambda}_1)\, p(x_\mathrm{m}|\boldsymbol{\lambda}_0,\boldsymbol{\lambda}_1,\epsilon,B)}{\int d\boldsymbol{\lambda}\, p(\boldsymbol{\lambda})\, p(x_\mathrm{m}|\boldsymbol{\lambda},\epsilon,B)}\,. \tag{8}$$

In Section IV we consider this process of marginalization over unwanted, or nuisance, parameters in more detail.

To calculate an upper limit, or rate interval, at a given confidence level $\mathcal{C}$, one integrates Eq. (8) over a region $V(\boldsymbol{\lambda}_0)$ such that

$$\mathcal{C} = \int_{V(\boldsymbol{\lambda}_0)} p(\boldsymbol{\lambda}_0|x_\mathrm{m},\epsilon,B)\,. \tag{9}$$

In general, the difficult part is selecting the region $V(\boldsymbol{\lambda}_0)$, especially in more than one dimension. There are several ways to do this: for example, one could marginalize over all but one of the parameters thus reducing the problem to a 1-d integral; or select the smallest volume $V(\boldsymbol{\lambda}_0)$ that gives the required probability. (This is sometimes called a highest posterior density interval [12].) In Sec. V, we investigate the properties of this type of rate interval based on the loudest event method.

III. UNKNOWN RATE AMPLITUDE

We have obtained the general expression for the posterior probability distribution of the parameters $\boldsymbol{\lambda}$ governing an astrophysical model based on an observed loudest event. In practice, the details of obtaining either a rate upper limit or a confidence interval on the model parameters will depend upon the details of the astrophysical model and its dependence upon the variables $\boldsymbol{\lambda}$. In this section, we simplify to the situation where the rate is dependent upon a single parameter μ which acts as an overall unknown amplitude, so that

$$\mathcal{R}(\mathcal{A}|\mu) = \mu\,\mathcal{R}_o(\mathcal{A})\,. \tag{10}$$

There are many instances where this simplification is well motivated physically. For example, in the search for coalescing binaries, it is typical to assume that the rate of binary coalescence is directly proportional to the blue light luminosity, while the constant of proportionality is unknown. Thus, the rate $\mathcal{R}_o$ will depend upon the distribution of blue light in the universe which can, in principle, be measured. Then, an observation can be used to set a limit on the rate μ. Similarly, in a search for other astrophysical sources, one might consider a population which is uniformly distributed in space, $\mathcal{R}_o(D) = \mu D^2$, where D is the distance.

We can use this form of the rate to simplify the general expression for the posterior. To begin, we introduce the quantity

$$R_o(x) = \int d\mathcal{A}\,\epsilon(x,\mathcal{A})\,\mathcal{R}_o(\mathcal{A})\,. \tag{11}$$

Then $R(x) = \mu R_o(x)$, and (at least in principle) $R_o(x)$ is known. In addition, we introduce the quantity $\Lambda(x)$, defined as

$$\Lambda(x) = \frac{|R_o'(x)|}{P_b'(x)}\left[\frac{R_o(x)}{P_b(x)}\right]^{-1}\,. \tag{12}$$

The quantity $\Lambda(x_\mathrm{m})$ encodes the likelihood that the loudest event is due to the foreground. In particular, in the limit that $\Lambda(x_\mathrm{m}) \to 0$, the event is definitely from the background. Alternatively, in the limit that $\Lambda(x_\mathrm{m}) \to \infty$, the loudest event is definitely from the foreground. We note that it depends upon the physical rate distribution $R_o(x_\mathrm{m})$ and its derivative $R_o'(x_\mathrm{m})$. Similarly, it depends upon the background distribution through $P_b(x_\mathrm{m})$ and it's derivative.

For an unknown rate amplitude, the above expressions can be substituted into Eq. (7) to obtain the posterior distribution

$$p(\mu|x_\mathrm{m},R_o,B) \propto p(\mu)\left[1+\mu\,\Lambda(x_\mathrm{m})R_o(x_\mathrm{m})\right]e^{-\mu R_o(x_\mathrm{m})}\,. \tag{13}$$

It is necessary to include the contribution from the background; if it is ignored, then $\Lambda \to \infty$ and the posterior distribution will generically be peaked away from zero, and go to zero for a zero rate. Thus, even if there is no evidence for a signal in the data, the posterior distribution for the rate will be inconsistent with zero events . In the next subsections, we will explicitly evaluate the posterior for various choices of the prior $p(\mu)$ and obtain rate limits.

The loudest event prescription can be applied to any form of background, provided the required quantities in Eq. (12) can be measured or estimated. In many experiments, one might expect the background events above a statistic value x to be Poisson distributed, with a rate $\nu(x)$ where ν is a non-increasing function of x. Then, it follows directly that

$$\begin{aligned} P_b(x) &= e^{-\nu(x)} \\ p_b &= \nu'(x)|e^{-\nu(x)}| \\ p_b(x)/P_b(x) &= |\nu'(x)|\,. \end{aligned} \tag{14}$$

We will make use of this to simplify examples. However, it is important to note that the loudest event statistic does not require the assumption of Poisson background.

A. Uniform Prior

The posterior distribution given in Eq. (13) can now be evaluated in some particular cases. This requires a choice of prior distribution. For simplicity, we begin with a uniform prior,

$$p(\mu) = \mathrm{const}\,. \tag{15}$$

This distribution is not normalizable. However, we can introduce a cutoff at large μ (well above the rate being probed by the given experiment) in order to render it normalizable. Physically, this is a reasonable choice of prior if there is no information available about the expected value of μ.

It is straightforward to differentiate Eq. (13) and see that its peak will be away from zero if and only if $\Lambda(x_\mathrm{m}) > 1$. That is, the peak will be away from zero provided it is more likely that the loudest event is from the foreground than from the background. If this is the case, then one might take this as an indication of a non-zero rate. The extent to which this is true is explored in Sec. V.

We integrate Eq. (13) to obtain an upper limit at confidence level $\mathcal{C}$ by solving

$$1 - \mathcal{C} = e^{-\mu R_o(x_\mathrm{m})} \left[1 + \left(\frac{\Lambda(x_\mathrm{m})}{1+\Lambda(x_\mathrm{m})} \right) \mu R_o(x_\mathrm{m}) \right] \quad (16)$$

for μ. It has been shown in [8] that setting the background to zero yields a conservative rate limit. In the Bayesian analysis, however, this yields a posterior probability distribution function which is peaked away from zero, and goes to zero at zero rate. This is clearly seen in Fig. 1 which shows the posterior distribution for three values of Λ including $\Lambda \to \infty$. This is not surprising as we have neglected the background, in which case the existence of a loudest event implies a non-zero rate. Although this does not invalidate the upper limit (indeed, it has been shown that the no background limit is conservative), it does mean that the posterior would not serve as a suitable prior for a future experiment, as it is inconsistent with a zero rate. Nevertheless, it is still possible to obtain the upper limit as

$$\mu_{90\%} = \frac{3.9}{R_o(x_\mathrm{m})} \,. \quad (17)$$

Similarly, the no-foreground limit can be obtained by taking $\Lambda = 0$. In this case, the 90% confidence limit tends to

$$\mu_{90\%} = \frac{2.3}{R_o(x_\mathrm{m})} \,. \quad (18)$$

Finally, we can consider the case where the loudest event is equally likely to be due to foreground or background, $\Lambda = 1$. In this case, we have,

$$\mu_{90\%} = \frac{3.3}{R_o(x_\mathrm{m})} \,. \quad (19)$$

The posterior distribution for the rate for these three possibilities is shown in Fig. 1.

B. Alternative Choices of Prior

An alternative distribution on the rate is the Jeffrey's prior, given by $p(\mu) \propto 1/\mu$. Like the uniform prior, it is not normalizable. This can be fixed by imposing a cutoff at a small value of x, but doing so artificially introduces a length scale into the problem. We can again compute the posterior distribution as

$$p(\mu|x_\mathrm{m}, R_o, B) \propto \left[\frac{1}{\mu} + \Lambda(x_\mathrm{m}) R_o(x_\mathrm{m}) \right] e^{-\mu R_o(x_\mathrm{m})} \,. \quad (20)$$

In the no-background case, this fixes the issue of the distribution being peaked away from zero. However, for the general case where the background is included, this distribution is not satisfactory due to the $1/\mu$ term, which will again need to be cut off at some value.

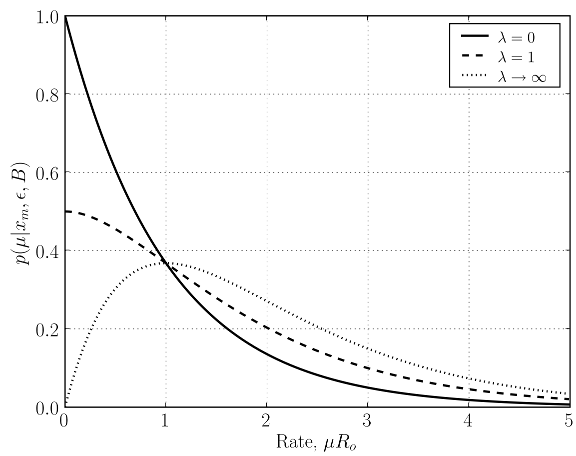

FIG. 1: The posterior probability density function $p(\mu|x_m, R_o, B)$ on the rate μ, assuming a uniform prior. The three curves correspond to three different values of the likelihood: a) $\Lambda = 0$ (solid line), the loudest event is definitely background and the distribution is exponential; b) $\Lambda = 1$ (dashed line), the loudest event is equally likely to be from the foreground or background, the distribution peaks at zero but the derivative vanishes there; c) $\Lambda \to \infty$ (dotted line), the loudest event is definitely from the foreground, the distribution is peaked away from zero.

Another alternative is the exponential prior. We first note that this arises naturally when a previous search has been completed. In this case, it is natural to use the posterior from a previous search as the prior distribution for a subsequent one. If the first search was performed using a uniform prior, the posterior is given by (13). Furthermore, in the event that the loudest event is most likely background $\Lambda \ll 1$. Then, we can conservatively rewrite the posterior as

$$p_c(\mu|x_\mathrm{m}, R_o, B) \propto e^{-\mu\Lambda(x_\mathrm{m})R_o(x_\mathrm{m})} \quad (21)$$

where we have made use of the fact that

$$1 + \mu\Lambda R_o \leq e^{\mu\Lambda R_o} \,. \quad (22)$$

It is straightforward to show that the rate limit at a given confidence level $\mathcal{C}$ inferred using this posterior is necessarily larger than that obtained using the original distribution. In this sense, the alternative distribution is conservative and the distribution has been cast as an exponential.

Starting with the exponential prior,

$$p(\mu) \propto e^{-r_o \mu} \,, \quad (23)$$

whose decay is governed by the quantity r_o, we can obtain a posterior distribution. It is beneficial to re-define Λ as

$$\Lambda(x) = \frac{|R_o'(x)|}{P_b'(x)} \left[\frac{R_o(x) + r_o}{P_b(x)} \right]^{-1} \,. \quad (24)$$

Then, the posterior distribution is given by

$$p(\mu|x_{\mathrm{m}}, R_o, B) \propto [1 + \mu\Lambda(x_{\mathrm{m}})(R_o(x_{\mathrm{m}}) + r_o)]\, e^{-\mu(R_o(x_{\mathrm{m}})+r_o)} \,. \tag{25}$$

As before, the posterior distribution is peaked away from zero if $\Lambda > 1$. In addition, the distribution is identical to that obtained using a uniform prior, only now we have effectively searched over $R_o(x_{\mathrm{m}}) + r_o$.

C. Comparison with Fixed Thresholds

Let us compare the loudest event statistic against a fixed threshold approach. In order to do this, we work with the example discussed in the introduction: $\nu(x) = e^{(8^2-x^2)/2}$, $R_o(x) = (8/x)^3$. The normalizations of these functions are chosen for simplicity so that $\nu(8) = R_o(8) = 1$, i.e. we expect one event at or above $x = 8$ and the rate is unity there. The value of the upper limit as a function of the actual loudest event is shown in Fig. 2a. The upper limit transitions smoothly from the zero foreground limit (at low values of x) to zero background limit (at large values of x). Figure 2b shows the distribution $p_b(x)$. This corresponds to the expected distribution of for the loudest event if it is due to the background. Then, by multiplying the upper limit by the expected distribution for the loudest event and integrating, we obtain the expected upper limit. In this example it is 2.64.

For comparison, the upper limit for a fixed threshold is presented in Figure 3. When calculating the upper limit for a fixed threshold, one simply counts the number of events n above the chosen threshold x_t and obtains a limit

$$\mu_{90\%} = \frac{F(n)}{R_o(x_t)} \tag{26}$$

where $F(n)$ is a known function for each integer n (see, for example, [13] for more details). In particular, when zero events are observed above the threshold, $F(0) = 2.3$. When performing a fixed threshold search, it is possible to take into account the expected background and, much as for the loudest event, neglecting to do so will lead to a conservative result. In Figure 3, we show the expected upper limit as a function of the threshold.

Clearly, in this example, the loudest event statistic is preferable to a fixed threshold, as it will provide a better expected upper limit value than the fixed threshold for *any* value of the threshold (with or without the background). We note that this result is specific to the details of the example under consideration; the key feature is that the background rate is a very steep function of x. Indeed, in [8], the same example was considered, but with an expected background of unity at $x = 4.5$ rather than $x = 8$, leading to a small range of values where the fixed threshold does beat the loudest event. However, as emphasized in that paper the attraction of the loudest event is that it is unnecessary to fix a threshold ahead of performing the search — the search itself determines the threshold.

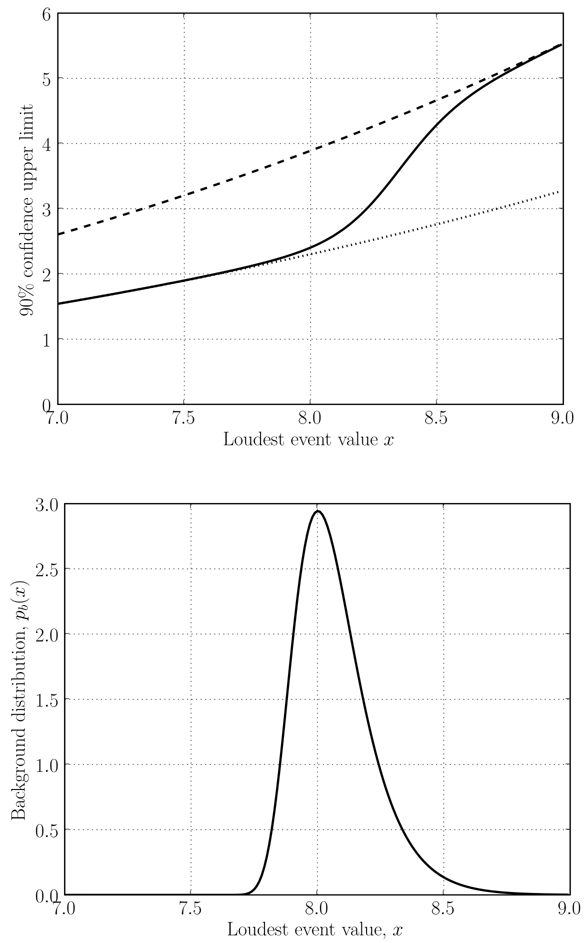

FIG. 2: a) The upper limit as a function of the observed loudest event. The solid line shows the value of the upper limit as a function of x. The dotted and dashed lines are given by $2.3/R_o(x)$ and $3.9/R_o(x)$. We see that the upper limit transitions smoothly from one to the other. At low values of x, the loudest event is very much consistent with the background, $\Lambda \approx 0$ and the upper limit is close to the dotted line. For larger values of x the loudest event is more consistent with foreground, $\Lambda \to \infty$, and the rate is more consistent with the dashed line. b) The probability distribution for the loudest event assuming that it is drawn from the background distribution, $p_b(x)$. Multiplying the upper limit curve by this distribution and integrating over x gives the expected value of the upper limit if the loudest event is from the background.

IV. MARGINALIZATION OVER UNCERTAINTIES

The rate of events in the data, $R(x|\boldsymbol{\lambda})$ in Eq. (3), is dependent upon the frequency of events and their amplitude distribution as well as the sensitivity of the search which is performed. In many cases, neither of these quantities will be precisely known. For example, the efficiency of an experiment is often measured via Monte-Carlo methods and therefore suffers from uncertainties due to the finite number of trials. If we expand our understanding of the parameters $\boldsymbol{\lambda}$ to further parametrize the uncertainties that can arise in the underly-

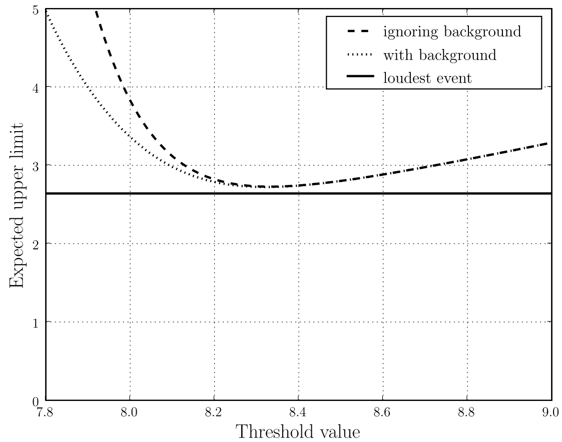

FIG. 3: Figure showing the expected upper limit as a function of the fixed threshold. The dashed line shows the upper limit obtained when ignoring the background, while the dotted line includes the background contribution. For large values of the threshold where the expected background is small, both limits approach $2.3/R_o(x)$ as expected. For low values of x, there is a good chance of many events above threshold which leads to a worse upper limit. The balance occurs at around a threshold value of $x = 8.3$. For reference, we also plot a horizontal line showing the expected upper limit from the loudest event. Interestingly, the loudest event will, on average, outperform the fixed threshold for any value of the threshold.

ing models and in measurements of efficiency, it is natural to marginalize over these uncertainties before computing an upper limit or rate interval. Just as the marginalization over uninteresting physical parameters (given in Eq. (8)) requires a prior distribution to be specified, the same is true of the uncertainties. This prior distribution would typically reflect the systematic and statistical errors estimate for the experiment.

As a particular example, consider the problem of the unknown rate amplitude presented in Sec. III and assume there is some uncertainty associated with the value of R_o. Typically, one might choose the prior to be a normal distribution peaked around the estimate value of $R_o(x_m)$. It is, however, unphysical for the rate to be zero, so the distribution would need to be truncated. A more natural choice is a log-normal distribution, for which the logarithm of R_o would be normally distributed, thereby guaranteeing that R_o is positive.

Here, we choose to make use of the γ-distribution, primarily because it can be analytically integrated. The γ-distribution is similar in shape (for small standard deviation) to both the Gaussian and log-normal distributions and in addition takes only non-negative values. For integer values k, the γ-distribution is given by

$$p(R_o|k,\theta) = \frac{R_o^{(k-1)}\, e^{-R_o/\theta}}{\theta^k\, k!}\,. \tag{27}$$

The mean is $\hat{R}_o = k\theta$ while the standard deviation is $\sigma_{R_o} = \sqrt{k}\theta$. Therefore, fractional standard deviation $\sigma_{R_o}/\hat{R}_o = 1/\sqrt{k}$, which tends to zero in the limit as $k \to \infty$, whereby we expect to recover the unmarginalized results.

The marginalized distribution is calculated by integrating over R_o,

$$p(\mu|x_{\rm m}, \hat{R}_o, k, B) = \int dR_o\, p(R_o|\hat{R}_o, k)\, p(\mu|x_{\rm m}, R_o, B)\,. \tag{28}$$

Making use of the distribution (13) for the rate and the expression for the γ-distribution given above, we obtain the marginalized distribution

$$p(\mu|x_{\rm m}, \hat{R}_o, k, B) = \frac{\hat{R}_o}{(1+\Lambda)}\left[\frac{1}{\left(1+\mu\hat{R}_o/k\right)^{k+1}} + \frac{\mu\Lambda\hat{R}_o(1+1/k)}{\left(1+\mu\hat{R}_o/k\right)^{k+2}}\right]\,. \tag{29}$$

In the limit that $k \to \infty$, we recover the previous distribution for μ as expected.

In order to examine the effect of marginalization, in Figure 4 we plot the unmarginalized posterior distribution for $\Lambda = 10$ along with three distributions obtained by marginalizing over different size systematic errors or uncertainties. These distributions are obtained from (29) with values of $k = 100, 16$ and 4 corresponding to errors of $10, 25$ and 50% respectively. As the systematic error increases, the posterior distribution for the rate gets broader; the value of the probability density function increases for large values of the rate. This causes an increase in the upper limit. Without taking into account any uncertainties, the 90% confidence upper limit is $3.8/R_o(x_{\rm m})$. For 10% systematic error, this increases only slightly to $3.9/R_o(x_{\rm m})$ while for 25 and 50% this increases further to $4.2/R_o(x_{\rm m})$ and $5.5/R_o(x_{\rm m})$ respectively. In Figure 5 we plot the upper limit as a function of the systematic error for four different values of Λ. The results are qualitatively similar to what was seen before — marginalizing over uncertainties will increase the upper limit and the larger the errors, the larger the effect.

A. Marginalization over uncertainties in Λ

In many cases, there will also be uncertainties in the precise value of Λ. These can be marginalized over in the same way as described above. Since the Λ dependence of the distribution (13) is straightforward, this can be done explicitly. For

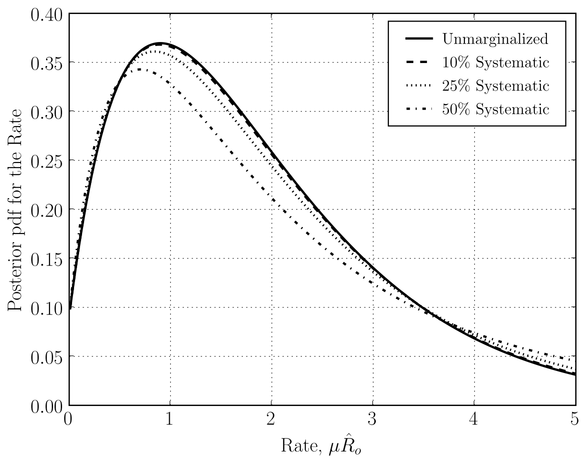

FIG. 4: The posterior probability density function on the rate for different sizes of systematic error. The curves were generated assuming a uniform prior and using $\Lambda = 10$. The solid line corresponds to the unmarginalized probability density function. The dot-dashed line gives the distribution marginalized over a 10% systematic uncertainty (equivalently $k = 100$ for the γ-distribution). With this level of uncertainty, the marginalized distribution is barely changed from the original. The dotted and dashed lines show the posterior for 25% ($k = 16$) and 50% ($k = 4$) systematic errors. As the systematic error increases the distribution broadens and consequently the upper limit increases.

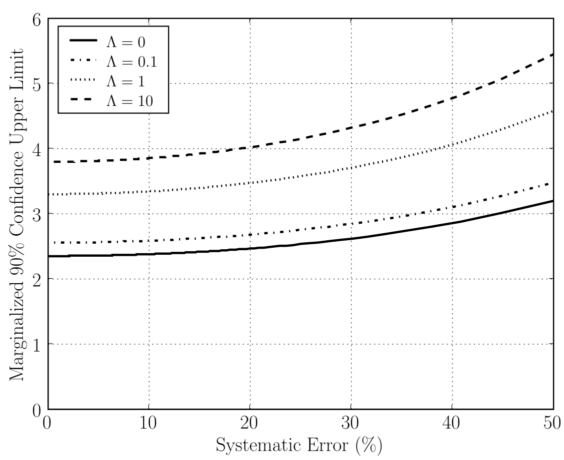

FIG. 5: The 90% confidence upper limit versus the size of the systematic error which is marginalized over (equivalent to $1/\sqrt{k}$ in the γ-distribution discussed in the text). The limit plotted for four different values of Λ: $0, 0.1, 1, 10$. In all cases, the upper limit increases with larger systematic error.

concreteness, let us take a uniform prior, in which case, the posterior distribution is:

$$p(\mu|x_\mathrm{m},\Lambda) = R_o(x_\mathrm{m})\frac{(1+\mu R_o(x_\mathrm{m})\Lambda)}{(1+\Lambda)}e^{-\mu R_o x_\mathrm{m}(x_\mathrm{m})} \quad (30)$$

Then, given a probability distribution $p(\Lambda)$, the marginalized distribution is

$$p(\mu|x_\mathrm{m}) = \int d\Lambda\, p(\Lambda)\, p(\mu|x_\mathrm{m}, R_o, \Lambda) \quad (31)$$

In this case, the above integral of is straightforward. Specifically, let us define

$$\xi = \int d\Lambda\, p(\Lambda)\, \frac{\Lambda}{(1+\Lambda)}\,. \quad (32)$$

Then, the posterior distribution following marginalization over Λ is given by

$$p(\mu|x_\mathrm{m},\Lambda) = R_o(x_\mathrm{m})\,[(1-\xi)+\mu R_o(x_\mathrm{m})\xi]\, e^{-\mu R_o x_\mathrm{m}(x_\mathrm{m})} \quad (33)$$

Suppose that Λ is distributed with expectation value $\hat{\Lambda}$ and variance σ_Λ^2. Then, to leading order,

$$\xi \approx \left(\frac{\hat{\Lambda}}{1+\hat{\Lambda}}\right) - \left(\frac{\sigma_\Lambda^2}{(1+\hat{\Lambda})^3}\right)\,. \quad (34)$$

From this, we notice two things. Firstly, even if the fractional uncertainties in Λ are of order unity, when $\Lambda \gg 1$ or $\Lambda \ll 1$, the second term is small compared to the first and can be ignored. Secondly, marginalizing over Λ only serves to decrease the value of ξ relative to the unmarginalized case. This is equivalent to reducing the likelihood that the loudest event is foreground and consequently will reduce the upper limit. Therefore, it is possible to neglect the marginalization of Λ as this is a conservative thing to do.

V. RATE INTERVALS

In Sec. III, we derived the upper limit on the rate μ based on the loudest event. However, in the case where the likelihood Λ of the event being foreground is large, we may prefer to give a rate interval rather than an upper limit. For a uniform prior, the mode μ_p of the posterior distribution for the rate (given in Eq. (13)) is non-zero whenever $\Lambda > 1$. Furthermore, in this case,

$$\mu_p = \frac{1}{R_o(x_\mathrm{m})} - \frac{1}{\Lambda R_o(x_\mathrm{m})}\,, \quad (35)$$

which asymptotes to $1/R_o$ for large values of Λ as one might expect. How significant an indicator of a non-zero rate is having the peak of rate distribution be non-zero? In order to examine this idea more precisely, we will describe a method of constructing a rate interval using the loudest event statistic.

At some confidence level $\mathcal{C}$, a rate interval is given by $[\mu_1, \mu_2]$ such that

$$\int_{\mu_1}^{\mu_2} p(\mu|x_m, \epsilon, B)\, p(\mu)\, d\mu = \mathcal{C}\,. \quad (36)$$

A supplementary condition is required to select a unique interval: we identify the interval which minimizes $|\mu_2 - \mu_1|$ and contains the mode of the distribution (or zero for $\Lambda < 1$).

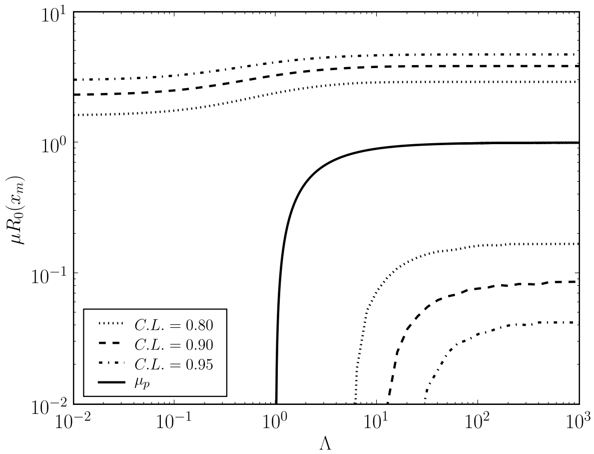

FIG. 6: The graph shows the behavior of the lower and upper boundaries of the rate interval, μ_1 and μ_2 respectively, as a function of the likelihood Λ. They are plotted for three different values of the confidence level $\mathcal{C}$ of 80%, 90% and 95%. The peak μ_p (solid line) approaches zero as Λ approaches one. As $\Lambda \to 0$, μ_2 agrees with no foreground upper limit treated above.

This condition clearly results in $\mu_1 = 0$ for small values of Λ, i.e. when the loudest event was likely to have arisen from the background, the rate interval on the process we wish to constrain includes zero rate.

For the uniform prior considered in Sec. III A, the dependence of μ_1, μ_2 and μ_p on Λ are shown in Fig. 6. For $\Lambda < 1$, $\mu_p = 0$ and consequently $\mu_1 = 0$, as expected. However, for a significant range of $\Lambda > 1$, even though the rate distribution is peaked away from zero, $\mu_1 = 0$ indicating that (at the given confidence) the rate interval still includes zero.

We can determine the precise value of Λ at which μ_1 becomes non-zero. For fixed Λ and $R_o(x_\mathrm{m})$, Eq. (36) gives μ_2 implicitly as a function of μ_1. The minimal interval condition is then just

$$\frac{d(\mu_2 - \mu_1)}{d\mu_1} = 0 \, . \tag{37}$$

Substituting $\mu_1 = 0$ into Eq. (36) and its derivative gives two equations which depend on μ_2 and Λ. As an example, consider a 90% confidence interval. In this case, μ_1 becomes non-zero, and the interval is bounded away from the origin, at value of $\Lambda \simeq 11.56$. This corresponds to $\mu_2 \simeq 3.807/R_o(x_\mathrm{m})$. This result is in good agreement with the values obtained numerically in Figure 6. In this context, it is interesting to return to Figure 4 which shows the posterior distribution for $\Lambda = 10$. Although this distribution is peaked well away from zero, the 90% confidence interval still includes zero.

VI. COMBINING RESULTS FROM MULTIPLE EXPERIMENTS

When performing a series of experiments, there is a very natural way to combine the results in a Bayesian manner. As discussed above, the calculation of a Bayesian upper limit requires the specification of a prior probability distribution for the rate μ. When a previous experiment has been performed, it is natural to use the posterior from the first experiment as the prior for the second. It is straightforward to show that the results are independent of the order of the experiments. (This does not depend upon the loudest event, rather it is a general Bayesian result). Begin by recalling that

$$p(\mu|x_1) = \frac{p(\mu)\, p(x_1|\mu)}{\int d\mu\, p(\mu)\, p(x_1|\mu)}. \tag{38}$$

For the second search, simply use $p(\mu|x_1)$ as the prior to obtain the posterior distribution on μ given the observations in both the first and second experiments:

$$p(\mu|x_1, x_2) = \frac{p(\mu)\, p(x_1|\mu)\, p(x_2|\mu)}{\int d\mu\, p(\mu)\, p(x_1|\mu)\, p(x_2|\mu)} \tag{39}$$

This is clearly symmetric in x_1 and x_2. It is straightforward to see that marginalization over nuisance parameters preserves this symmetry.

Next, let us consider the effect of taking a single search and splitting it into two halves, which can be combined to produce an upper limit in the manner described above. Naively, it appears that splitting the search will give a lower rate limit, since we will be using a "quieter" loudest event for half the search. If this were the case, then it would seem that splitting the search into ever shorter searches would improve the upper limit indefinitely. As we shall see, the result is not so clear cut, and it depends critically upon the foreground and background distributions $R_o(x)$ and $P_b(x)$.

Consider an experiment performed for some given time T, and assume that both the foreground and background rates are constant over time. We would then like to compare the (expected) upper limit from the full search to that obtained by splitting the data in two parts of length T_1 and T_2 and calculating a combined upper limit from the two searches. Let us assume, without loss of generality, that the loudest event overall in the search occurs in the first half of the search with a statistic value of x_1, and the loudest event in the second half of the search has a statistic value x_2. Then, we can calculate the upper limit from the search (taking it as a single entity) and from the split search.

The posterior from the single search is given by

$$p(\mu|x_1) \propto p(\mu)\, (1 + \mu R_o(x_1)\Lambda(x_1))\, e^{-\mu R_o(x_1)} \, . \tag{40}$$

While for the split search, the posterior for each part is given by given by

$$p(\mu|x_i) \propto p(\mu)\, [1 + \mu\Lambda(x_i) R_o(x_i)]\, e^{-\mu R_o(x_i) T_i/T} \tag{41}$$

where $i = 1, 2$ label the two parts of the search. Let $\alpha = T_1/T$, then the combined posterior distribution for the split search is

$$p(\mu|x_1, x_2) \propto p(\mu)\,[1 + \mu\Lambda(x_1)R_o(x_1)]\,[1 + \mu\Lambda(x_2)R_o(x_2)]\,e^{-\mu[\alpha R_o(x_1)+(1-\alpha)R_o(x_2)]}. \quad (42)$$

Notice, for the split search, the exponential decay term is at least as large as for the single search, with equality only if $x_1 = x_2$. This tends to make the upper limit obtained in the split search smaller than that of the single search. In contrast, the polynomial prefactor is always more significant for the split search (i.e. it grows more steeply with μ). This tends to make the upper limit larger. Whatever the form of $\Lambda(x)$ and $R_o(x)$, it is clear that in the case where $x_1 = x_2$, splitting the search will lead to a larger limit. Meanwhile if $x_2 \ll x_1$, the split search will give a numerical smaller limit.

To quantify these statements, we perform a detailed analysis for the uniform prior. When $\Lambda \ll 1$, the posterior distribution for the single search can be approximated as

$$p(\mu|x_1) \simeq R_o(x_1)[1 - \Lambda(x_1)]e^{-\mu R_o(x_1)[1-\Lambda(x_1)]} \quad (43)$$

while the posterior for the split search becomes

$$p(\mu|x_1, x_2) \simeq c(x_1, x_2)e^{-\mu c(x_1,x_2)} \quad (44)$$

where

$$c(x_1, x_2) = R_o(x_1)[\alpha - \Lambda(x_1)] + R_o(x_2)[(1 - \alpha) - \Lambda(x_2)]\ . \quad (45)$$

Within the context of this approximation, it is then easy to write down the upper limit for each distribution. In particular,

$$\mu_{\rm single} = \frac{\ln(1.0 - \mathcal{C})}{R_o(x_1)[1 - \Lambda(x_1)]} \quad (46)$$

for the single search; for the split search

$$\mu_{\rm split} = \frac{\ln(1.0 - \mathcal{C})}{c(x_1, x_2)}\ . \quad (47)$$

Hence, the single search will give a smaller upper limit if

$$\Lambda(x_2) > (1 - \alpha)\left[1 - \frac{R_o(x_1)}{R_o(x_2)}\right] \quad (48)$$

For comparison, the ratio of upper limits $\mu_{\rm single}/\mu_{\rm split}$ are plotted versus $\Lambda(x_2)$ for several different values of $R_o(x_1)/R_o(x_2)$ and $\alpha = 0.5$. For small $\Lambda(x_2)$, the single search gives a smaller upper limit when the condition in Eq. (48) is satisfied.

It is interesting to consider the example of Poisson distributed background, where $P_b(x, T_i) = e^{-\nu(x)T_i/T}$ to gain further insight into this result. If we assume that each of the halves of the search has the same $P_b(x, T/2)$, then the distribution of the loudest event x_1 is given by $P_b(x_1, T/2)^2 = P_b(x_1, T)$, while the distribution of the loudest event for the other half is given by $P_b(x_2)\,(2 - P_b(x_2))$. Then, we can easily obtain the median value for x_1 and x_2 as

$$\nu(x_1) \;=\; \ln(2) = 0.7 \quad (49)$$

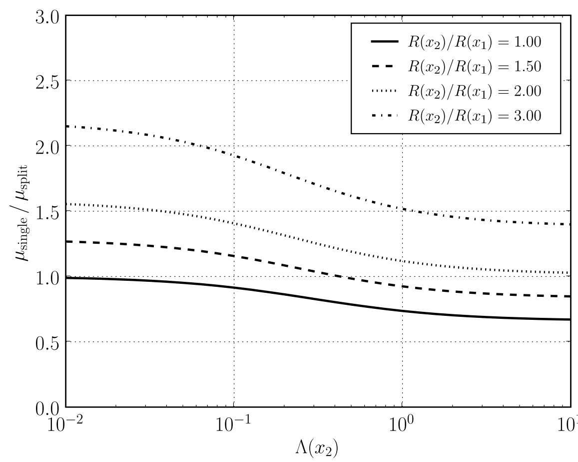

FIG. 7: The ratio of $\mu_{\rm single}$ to $\mu_{\rm split}$ as a function of $\Lambda(x_2)$ for several values of $R_o(x_2)/R_o(x_1)$. The figure was generated for $\Lambda(x_1) = 1.0$ and $\alpha = 0.5$. In general, there is only a weak dependence on this value; the curves steepen a little for smaller value of $\Lambda(x_1)$, but look qualitatively similar. Note also that for most sensible choices of amplitude statistic x, one expects $\Lambda(x_2) \le \Lambda(x_1)$. The plot is extended to $\Lambda(x_2) = 10$ for completeness.

and

$$\nu(x_2) \;=\; -2\ln(1 - \sqrt{2}/2) = 2.5 \quad (50)$$

Taking the difference of the above equations, we obtain

$$\nu(x_2) - \nu(x_1) = 1.8 \quad (51)$$

Then if we define $\Delta x = x_1 - x_2$, we have

$$\nu(x_1) \approx \nu(x_2) - \Delta x|\nu'(x_2)|. \quad (52)$$

Hence

$$\Delta x \approx \frac{1.8}{|\nu'(x_2)|}. \quad (53)$$

and similarly

$$R_o(x_2) - R_o(x_1) \approx \Delta x|R_o'(x_2)|\,. \quad (54)$$

For a Poisson background, from (12) and (14), the likelihood can be written as

$$\Lambda(x_2) = \frac{R_o'(x_2)}{R_o(x_2)|\nu'(x_2)|} \quad (55)$$

Then, substituting in our expressions for $|\nu'(x_2)|$ from (53) and $R_o'(x_2)$ from (54) yields

$$\Lambda(x_2) \approx \left(\frac{1}{1.8}\right)\left[1 - \frac{R_o(x_1)}{R_o(x_2)}\right]\ . \quad (56)$$

Recall that if the inequality in equation (48) is satisfied, then the single search is expected to give a smaller upper limit that the split search. From Eq. (56) above, we see that the inequality is indeed satisfied. Therefore, we conclude that the single search is expected to give a smaller limit on average in this approximation.

VII. DISCUSSION

The loudest event statistic is just one method of taking account of the quality of an event in the interpretation of a search. In this paper, we have presented further exploration of the method including the discussion of marginalization over uncertainties in the input model. The Bayesian approach allows simple accounting of these uncertainties by integrating them out.

In addition, we showed how the method could be used to determine a rate interval. Once again, this is not the most powerful method of determining an interval (in the sense that using more than one event would lead to a more strongly peaked distribution and, consequently, a narrower interval). Nevertheless, the approach shows that a rate interval arises when the likelihood that the event is signal becomes large enough.

Finally, we presented a discussion of combining the results from multiple searches to determine a single upper limit. It was shown that the limit obtained by combining two searches of equal duration is, in general, different to the limit obtained by performing a single search of equivalent duration. What conclusion to draw from this is unclear since the notion of better depends on the true value of the rate being explored.

Even though physicists have a deep appreciation for probabilistic phenomena in nature, it is often tempting to talk about better upper limits by using one method or another. This is, of course, a flawed approach. In fact, it is the experiment that one should choose not the statistical method. Nevertheless, some experiments may be more powerful than others. For example, it would be ill-conceived to use the loudest event method to determine a rate interval in an experiment which is likely (in the sense of prior probability) to generate more than one loud event that could be considered to arise from the phenomenon of interest. Indeed, these considerations lead back to an experiment more like the standard threshold approach.

Acknowledgments

We would like to acknowledge many useful discussions with members of the LIGO Scientific Collaboration inspiral analysis group which were critical in the formulation of the methods and results described in this paper. This work has been supported in part by NSF grants PHY-0200852 and PHY-0701817; PRB is grateful to the Research Corporation for support by a Cottrell Scholar Award; SF was funded in part by the Royal Society. 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-0107417. This paper has LIGO Document Number LIGO-P070076-00-Z.

[1] B. Allen et al., Phys. Rev. Lett. **83**, 1498 (1999).
[2] B. Abbott et al. (LIGO Scientific Collaboration), Phys. Rev. D **69**, 122001 (2004).
[3] B. Abbott et al. (LIGO Scientific Collaboration), Phys. Rev. D **72**, 082002 (2005), gr-qc/0505042.
[4] B. Abbott et al. (LIGO Scientific Collaboration), Phys. Rev. D **72**, 082001 (2005), gr-qc/0505041.
[5] B. Abbott et al. (LIGO Scientific Collaboration), Phys. Rev. D **73**, 062001 (2006).
[6] B. Abbott et al. (LIGO Scientific), Phys. Rev. **D69**, 082004 (2004), gr-qc/0308050.
[7] X. Siemens et al., Phys. Rev. **D73**, 105001 (2006), gr-qc/0603115.
[8] P. R. Brady, J. D. E. Creighton, and A. G. Wiseman, Class. Quant. Grav. **21**, S1775 (2004).
[9] R. Cousins, Nucl. Instrum. Meth. **A337**, 557 (1994).
[10] S. Yellin, Phys. Rev. **D66**, 032005 (2002), physics/0203002.
[11] P. Brady, S. Fairhurst, et al. (2007), grqc/0000.
[12] W.-M. Yao et al., J. Phys. G: Nucl. Part. Phys. **33**, 1 (2006).
[13] W. M. Yao et al. (Particle Data Group), J. Phys. **G33**, 1 (2006).