LASER INTERFEROMETER GRAVITATIONAL WAVE OBSERVATORY
- LIGO -
CALIFORNIA INSTITUTE OF TECHNOLOGY
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

Document Type	**LIGO-T030017-00-Z**	2003/02/23

Notes on Rate vs. Strength Upper Limits

Lee Samuel Finn

Distribution of this draft:

LIGO I Collaboration

California Institute of Technology
LIGO Project - MS 51-33
Pasadena CA 91125
Phone (626) 395-2129
Fax (626) 304-9834
E-mail: info@ligo.caltech.edu

Massachusetts Institute of Technology
LIGO Project - MS 20B-145
Cambridge, MA 01239
Phone (617) 253-4824
Fax (617) 253-7014
E-mail: info@ligo.mit.edu

WWW: http://www.ligo.caltech.edu/

Processed with LaTeX on 2003/02/23

Abstract

A principle result of the burst group S1 analysis is an upper limit on the rate of gravitational wave strain events vs. event strength. In these notes we try to give precise definition to this result and describe how, given that definition, it should calculated. The result calculated on the S1 data is only an approximation to the definition given here and we describe the approximations involved.

The analysis approach described here for event data has broader applicability the determination of rate v. strength curves. That broader applicability is the subject of a separate technical note.

$Id: T030017.tex,v 1.4 2003/02/23 22:21:06 lsf Exp $

Contents

1 Rate vs. Strength: a definition

A principle result of the burst group S1 analysis is an upper limit on the rate of gravitational wave "events" $\dot{n}_S$ vs. event "strength" $\vec{h}_0$. The graphical form of the result is a single curve in the plane of strength (ordinate) vs. rate (abscissa). The qualitative shape of the curve is $K = (\dot{n}_S - 1)(\vec{h}_0 - 1)$. What is the meaning of a point on this curve?

Here we take a point on the result curve to be the *bound* on number of events per unit time associated with a population of sources characterized by a strain $\vec{h}_0$. To obtain a point on the curve we must determine the likelihood associated with an observation of N strain events in a detector livetime T in the presence of a background and a possible signal foreground (characterized by $\vec{h}_0$) with an event rate $\dot{n}_S$.

Note that this meaning refers to a point on the curve, irrespective of other points on the curve. The distinction is subtle, but crucial. Each point on the curve is an upper limit associated with a given degree of confidence. The region excluded by the curve is not, however, associated with that same degree of confidence since it is a statement about a group of not entirely independent hypotheses considered separately on the same data.

We assume that at the end of the data processing pipeline we have a set of events characterized by strain amplitudes H_+ and $H_\times$ and wave propagation direction $\vec{m}$. The S1 burst group analysis pipeline does not generate all this information; nevertheless, the S1 burst group analysis can be described as an approximation to the analysis described here on these more detailed observations. The remainder of these notes describe how points on the result curve are defined in terms of the observations and a set of approximations that reduce this general analysis to the one used in S1.

2 Source population and gravitational wave strain events

In this section we define the source population and relate it to the population's contribution to the events identified in the analysis.

Consistent with the nature of our result as an excluded region in the $(\vec{h}_0, \dot{n}_S)$ we assume a set of population models parameterized by the source "strength" parameter $\vec{h}_0$ and an intrinsic source event rate $\dot{n}_S$,

$$\dot{n}_S = \begin{pmatrix} \text{rate of gravitational wave events from the} \\ \text{population characterized by } \vec{h}_0 \end{pmatrix} \tag{1}$$

We are concerned with bounding $\dot{n}_S$ for fixed $\vec{h}_0$.

Note that, more generally, $\vec{h}_0$ could be a set of parameters $\vec{h}_0$ describing a higher-dimensional set of source and source population models. Reflecting this more general nature of the analysis we will write $\vec{h}_0$ where applicable, though our prime focus in these notes is rate v. strength analysis results where $\vec{h}_0$ is the scalar source strength.

Focus attention on the particular model characterized by $\vec{h}_0$. This population leads to gravitational wave events incident on the detector array. The individual events are characterized at the very least by two wave polarization amplitudes $(h_+, h_\times)$ and the wave propagation direction $\vec{n}$. The population leads to a distribution of events at the detector that we write as $p(\vec{h}|\vec{h}_0)$:

$$p(\vec{h}|\vec{h}_0) = \begin{pmatrix} \text{probability of strain event of characterized} \\ \text{by } \vec{h} \text{ given population characterized by } \vec{h}_0 \end{pmatrix} \tag{2}$$

where

$$\vec{h} = \begin{pmatrix} \text{Parameters describing gravitational} \\ \text{waves incident on detector array} \end{pmatrix}. \tag{3}$$

We assume, at a minimum, that $\vec{h}$ includes wave polarization amplitudes, referred to a source model, and the propagation direction of the wavefront. It may also include other parameters that characterize the source population: e.g., the source population may be inhomogeneous with parameters describing the characteristics of the inhomogeneity, or it may consist of several discrete types of sources with a discrete parameter describing which source type the amplitudes h_+ and $h_\times$ are referred to. Note that, as a set of parameters, $\vec{h}$ is necessarily model dependent.

A particular event $\vec{h}$ incident on the detector may or may not lead to an observed event. Represent the probability of a real gravitational wave event leading to an observed event by ϵ:

$$\epsilon(\vec{h}) = \begin{pmatrix} \text{probability that the event } \vec{h} \\ \text{gives rise to a detector event} \end{pmatrix} \tag{4}$$

Note that depends both upon the source and source population model, the analysis that identifies events, and the noise character and the instrument calibration.[1] With these specified, however, $\epsilon(\vec{h})$ is readily determined by simulation.

Each observed event is characterized by at least a propagation direction $\vec{m}$ and an amplitude in each polarization $(H_+, H_\times)$. Let $\vec{H}$ represent the full description of the observed event:

$$\vec{H} = \begin{pmatrix} \text{Parameters describing gravitational} \\ \text{wave event identified in detector array} \end{pmatrix}. \tag{5}$$

Note that the dimensionality and details of the parameterization $\vec{H}$ depends on the character of the detector array (e.g., number of interferometers) and the nature of the analysis that identifies an event.

An important relationship is the one between actual events, described by $\vec{h}$ and detected events, described by $\vec{H}$. Let $q(\vec{H}|\vec{h})$ be the probability that the real event $\vec{h}$, *if observed*, leads to the characterized observation $\vec{H}$:

$$q(\vec{H}|\vec{h}) = \begin{pmatrix} \text{probability that observed event associated} \\ \text{with actual event } \vec{h} \text{ is characterized by } \vec{H} \end{pmatrix}. \tag{6}$$

The probability $q(\vec{H}|\vec{h})$ can be thought of as the uncertainty with which a detection can determine the character of the signal described by $\vec{h}$. Note that q depends on the source and source population model, the nature of the analysis, and the nature of the detector noise and calibration q. Once these are specified, however, q is readily determined, along with $\epsilon(\vec{h})$, by simulation.

The contribution of gravitational wave events associated with the population $\vec{h}_0$ to the detector output is thus described by the *foreground event distribution*

$$P_F(\vec{H}|\vec{h}_0) = \begin{pmatrix} \text{probability of observing the event described} \\ \text{by } \vec{H} \text{ from the population characterized by } \vec{h}_0 \end{pmatrix} \tag{7}$$

$$\propto \int d^n h \, q(\vec{H}|\vec{h})\epsilon(\vec{h})p(\vec{h}|\vec{h}_0) \tag{8}$$

where $d^n h$ is the measure on $\vec{h}$. For example, if $\vec{h}$ is the polarization amplitudes h_+ and $h_\times$ and the wave propagation direction $\vec{n}$ then $d^n h$ is $dh_+ dh_\times d^2S$, where d^2S is the surface element on the sphere described by the wave propagation direction. Note that $P_F(\vec{H}|\vec{h}_0)$ is completely described by the population model $p(\vec{h}|\vec{h}_0)$ and the analysis pipeline as characterized by simulation.

Associated with the source population is a total event rate $\dot{n}_S$. Not every source event leads to an observed event. The fraction of source events that lead to observed events is the total detection efficiency, which depends on the source population model, the detector noise and calibration, and the analysis methodology that identifies gravitational wave events. Writing the efficiency as $\epsilon(\vec{h}_0)$

[1] We assume here that the noise and instrument calibration are stationary. The problem of analysis in the presence of non-stationary noise and/or instrument calibration is described in section 6.1.

we have[2]

$$\begin{aligned} \dot{n}_F &= \left(\text{ rate of observed foreground events } \right) & (9) \\ &= \dot{n}_S \epsilon(\vec{h}_0) & (10) \\ \epsilon(\vec{h}_0) &= \int d^n h \, \epsilon(\vec{h}) p(\vec{h}|\vec{h}_0) & (11) \end{aligned}$$

3 Background distribution and event rate

Observed events may arise from the source population, in which case they are drawn from the distribution $P_F(\vec{H}|\vec{h}_0)$, or from environmental or instrumental artifacts. We refer to the distribution of events associated with environmental or instrumental artifacts as the background distribution

$$P_B(h)dh = \left(\begin{array}{l} \text{fraction of background} \\ \text{events in interval } [h, h+dh] \end{array} \right). \tag{12}$$

In addition to the background distribution, time-delay analysis also determines the expected rate of background events:

$$\dot{n}_B = \left(\begin{array}{l} \text{expected rate of background} \\ \text{events of any amplitude} \end{array} \right). \tag{13}$$

The background distribution and its rate may be estimated from time-delay coincidence analysis assuming that there is no preference for "zero-delay" background disturbances in the gravitational wave channel that cannot be vetoed by other means.

4 The likelihood function

The likelihood is the probability of a particular observation under a fixed hypothesis. In our case the hypothesis is that there is a source population characterized by "strength" $\vec{h}_0$ and an event rate $\dot{n}_S$ and our observation is a set of N observed events $\vec{H}$:

$$\left\{ \vec{H} \right\} = \left\{ \vec{H}_n : n = 1 \ldots N \right\}. \tag{14}$$

Focus first on the probability of a single event $\vec{H}$. That event may be foreground or background. The rate of foreground events is the product of the detection efficiency $\epsilon(\vec{h}_0)$ (cf. eqn. 11) and the signal event rate $\dot{n}_S$, which is what we wish to determine. Write the foreground event rate in terms of the background event rate $\dot{n}_B$ and a parameter α, $\alpha \in [0, 1)$:

$$\epsilon(\vec{h}_0)\dot{n}_S = \dot{n}_F = \dot{n}_B \frac{\alpha}{1-\alpha}. \tag{15}$$

As defined the parameter α is the probability that a particular event is a foreground event. In terms of α the probability of a particular event $\vec{H}$ is thus

$$P(\vec{H}|\vec{h}_0, \dot{n}_B, \dot{n}_S) = (1-\alpha) \, P_B(\vec{H}) + \alpha P_F(\vec{H}|\vec{h}_0). \tag{16}$$

[2] Again, we assume here stationary detector noise, calibration, etc., and treat the case of non-stationarity in section 6.1 below.

Now assume that gravitational wave events are independent of each other, and that the same is true of background events. The probability of making the observation $\left\{\vec{H}\right\}$ is then product of the probability of observing N events times, which is given by the Poisson distribution, and the probability that the N observed events are characterized by the particular $\vec{H}_n$, or

$$P\left(\{H\}\,|\vec{h}_0, T, \dot{n}_B, \dot{n}_S\right) = P(N|\mu)\begin{cases} 1 & N=0 \\ \prod_{n=1}^{N} P(\vec{H}_n|\vec{h}_0, \dot{n}_B, \dot{n}_S) & N>1 \end{cases} \tag{17}$$

where

$$P(N|\mu) = \frac{\mu^N}{N!}e^{-\mu} \tag{18}$$

is the Poisson distribution and

$$\mu = T\left(\dot{n}_B + \dot{n}_S\epsilon(\vec{h}_0)\right) \tag{19}$$

is the expected number of events in an observation of livetime T.

From the likelihood and the observation $\left\{\vec{H}\right\}$ we can find the bounds on $\dot{n}_S$ for fixed $\vec{h}_0$ by the usual techniques (e.g., Feldman & Cousins).

5 Approximations and the S1 analysis

In the S1 analysis the burst group made use only of the number of coincident events: neither the event amplitudes nor the propagation direction were determined or played a role in the analysis. In terms of the analysis described above this corresponds to the following set of approximations:

$$P_F(\vec{H}|\vec{h}_0) \propto 1, \tag{20}$$
$$P_B(h) \propto 1 \tag{21}$$

i.e., we do not distinguish between the relative likelihood of different foreground, or background, event amplitudes or incident directions. Under these two approximations the likelihood (cf. eqn. 17) is the Poisson distribution:

$$P(N|\vec{h}_0, \dot{n}_B, \dot{n}_S) = P(N|\mu) \tag{22}$$

where μ is given as before by the livetime T, the background rate $\dot{n}_B$ and signal rate $\dot{n}_S$ by equation 19:

$$\mu = T\left[\dot{n}_B + \epsilon(\vec{h}_0)\dot{n}_S\right]. \tag{23}$$

An observation of N events thus bounds $\left[\epsilon(\vec{h}_0)T\dot{n}_S\right]$ and determines a point on the result curve.

While it is gratifying that what was done in the S1 analysis can be related to a rigorous analysis, it must be emphasized that *these are very poor approximations.*

6 Discussion

6.1 Non-stationarity

The detector noise and calibration are not steady over the entire observation T. We can accommodate a time-varying noise and calibration if we can treat the noise and calibration as piecewise constant in time and know in what interval each of the N events in the observation $\left\{\vec{H}\right\}$ occurs.

Partition the total observation time T into M sub-intervals of duration t_k, $\sum_k^M t_k = T$, in which the noise and calibration are constant. Similarly partition the observation $\left\{\vec{H}\right\}$ into M disjoint sub-observations $\left\{\vec{H}\right\}_k$, with the union of the $\left\{\vec{H}\right\}_k$ equal to $\left\{\vec{H}\right\}$, such that all the events in $\left\{\vec{H}\right\}_k$ occur in the interval t_k. Associated with each sub-observation is the likelihood of making that observation given the expected background rate in the given interval: $P\left(\left\{\vec{H}\right\}_k |\vec{h}_0, t_k, \dot{n}_{B,k}, \dot{n}_S\right)$. *Note that the background event rate $\dot{n}_B$ and the distributions $P_B(\vec{H})$ and $P_F(\vec{H})$ will in general be different in each sub-interval.* The likelihood for the complete observation of duration T is then

$$P\left(\left\{\left\{\vec{H}_k\right\} : k = 1 \ldots M\right\} |\vec{h}_0, \{t_k, \dot{n}_{B,k}\}, \dot{n}_S\right) = \prod_{k=1}^{M} P(\left\{\vec{H}\right\}_k |\vec{h}_0, t_k, \dot{n}_{B,k}, \dot{n}_S). \tag{24}$$

From the likelihood we can derive the bound on $\dot{n}_S$ in the usual way.

Handling non-stationarity thus reduces to identifying epochs over which the noise and calibration are approximately stationary. Residual non-stationarity in each epoch will lead to a systematic error in the analysis. The degree to which stationarity should be required in an epoch is thus set by the level of the other systematic errors in the analysis.

Tracking calibration line amplitudes provides one method of identifying epochs over which the calibration is stationary. Observing the time dependent rate of background events and using a Bayesian Block analysis (cf. Scargle) is a possible approach to determining epochs when the noise is stationary.

6.2 Background rate uncertainty

The background rate $\dot{n}_B$ is determined experimentally. Associated with the experimental background rate is an uncertainty. Let

$$P_B\left(\dot{n}_B\right) d\dot{n}_B = \left(\text{ degree of belief that } \dot{n}_B \text{ is in } [\dot{n}_B, \dot{n}_B + d\dot{n}_B). \right) \tag{25}$$

We can marginalize the likelihood over this uncertainty, obtaining a new likelihood that is independent of uncertain $\dot{n}_B$

$$P(\left\{\vec{H}\right\} |\vec{h}_0, T, \dot{n}_S) = \int d\dot{n}_B P_B(\dot{n}_B) P(\left\{\vec{H}\right\} |\vec{h}_0, \dot{n}_B, \dot{n}_S) \tag{26}$$

The uncertainty $P_B\left(\dot{n}_B\right)$ may be estimated by making many estimates of the background rate, all at different delays, as long as the delays are much greater than any residual correlation time in the input time series from which the events are determined.

6.3 Improving the S1 analysis

The principal obstacle to improving the analysis undertaken in S1 is the determination of

- $\vec{H}$ for observed events;
- the background distribution $P_B(\vec{H})$;
- the foreground distribution $P_F(\vec{H}|\vec{h}_0)$;
- the intervals over which the noise and calibration are constant.

6.4 Goodness-of-fit

Any observation $\left\{\vec{H}\right\}$ will yield a bound on $\dot{n}_S$, even if the observation is, itself, very unlikely given our state of knowledge regarding the expected distribution of events in $\vec{H}$, background rate $\dot{n}_B$, and source population model $\vec{h}_0$. The value of the likelihood for the observation $\left\{\vec{H}\right\}$ provides a measure of the degree to which the observations are expected in the context of the model. Focus attention on the maximum of the likelihood over the source rate $\dot{n}_S$ given the observation. Simulations for this $\dot{n}_S$ will determine a distribution of observations and, correspondingly, values of the likelihood under the assumption that the rate is $\dot{n}_S$. The value of the likelihood for the actual observation can be compared to this distribution in order to determine how exceptional the observation is. If the observation is too exceptional given the best-fit (i.e., the maximum likelihood value of) $\dot{n}_S$ then we may wish to regard the bound on $\dot{n}_S$ as suspect.

6.5 Alternatives to rate vs. strength

Having determined the likelihood $P\left(\left\{\vec{H}\right\}|\vec{h}_0, T, \dot{n}_B, \dot{n}_S\right)$ we can proceed to find the *joint* bound on $(\vec{h}_0, \dot{n}_S)$: i.e., instead of finding the bound on the rate assuming the source distribution characterized by $\vec{h}_0$ we can find the region in $(\vec{h}_0, \dot{n}_S)$ space that best explains the observations.