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

Document Type	**LIGO-T030049-00-Z**	2003/03/31

Proposal for LIGO/TAMA Data Analysis

Lee Samuel Finn

Distribution of this draft:

LIGO/TAMA Analysis Team

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 2003/03/31

Abstract

We make a proposal for the detailed statistical approach to the joint analysis of the LIGO & TAMA data to bound the rate density of a galactic population of gravitational wave burst sources.

$Id: T030049.tex,v 1.7 2003/03/31 23:43:06 lsf Exp $

Contents

1 Introduction

LIGO and TAMA will analyze jointly data taken during the period 14 February to 14 April. A primary target of this joint data analysis effort is to bound the rate density of a galactic population of gravitational wave burst sources of unknown origin and, correspondingly, unknown waveform or other character. In these notes we describe a statistical approach to this problem based on the distributional properties of candidate gravitational wave events identified in each detector and characterized by a signal-to-noise ratio and time-of-arrival.

2 From source population to candidate gravitational wave events

2.1 Introduction and nomenclature

The beginning of our statistical analysis is a set of *candidate* gravitational wave events. Candidate events are each characterized by a set of properties (e.g., amplitude, arrival time, frequency, bandwidth, etc.). The set of candidate gravitational wave events consists of two disjoint subsets: one of events whose origin is noise, and which we refer to as *background* events, and the other of events whose origin is gravitational, and which we refer to as *foreground* events.

Gravitational wave bursts incident on our detector array arise from some astrophysical source and population distribution. Not all gravitational wave bursts arising from some source population and incident on our detector array are detected as candidate events. We refer to the set of all bursts events drawn from some population as *source events*.

Part of our goal is to determine, from the set of all candidate events, what fraction are foreground events. Knowing the relationship between foreground events and source events we can determine, from the number of foreground events, the number of source events. Another part of our goal is to bound the properties of a source and population distribution from which the observed foreground events can be thought to have derived based on the properties (amplitude, arrival time, frequency, bandwidth, etc.) of the candidate events and their distribution.

Gravitational wave bursts incident on a gravitational wave detector array can be characterized by their propagation direction and waveform in each of two polarization modes. Consistent with our focus on gravitational wave bursts of unknown origin we can't suppose we know the waveform or how the sources are distributed throughout space. Nevertheless, *we can interpret what we observe in our detector in terms of a source and population model*. The more detailed our characterization of the observed bursts and the greater their number the better we can distinguish between alternative models. Here we assume that candidate events are characterized by an overall event amplitude, though nothing we do in what follows restricts us from also considering candidate events that are characterized by separate amplitudes in each polarization mode, propagation direction, frequency, bandwidth, etc. We further assume that our source and population model has a single component distributed according to a fixed spatial distribution, in which case the model can be characterized by two parameters: an event rate $\dot{n}_S$ and an intrinsic source amplitude h_0 (though, again, nothing we do here restricts us from considering source and population models that have multiple components with different rates, more complicated luminosity functions, and multiple or parameterized spatial distributions.) For example, our source model may consist of axisymmetric sources that radiate a Gaussian pulse of fixed width and amplitude, and our population model may be a galactic distribution with a fixed disk scale height, bulge size and shape, etc. Our goal in this note to describe how we can relate the observation of a number of candidate events characterized by amplitude to the intrinsic amplitude h_0 and event rate $\dot{n}_S$ of our source population model.

In the first subsection below we describe the relationship between source events and an underlying source population model that we assume for the purpose of interpretation. In the following subsection we relate source events to foreground candidate events, including the relation between the source event characterization and the candidate event characterization. Background events are the subject of section 3. The likelihood for individual events and for an observation of N events is given in section 4 and Bayesian and Frequentist statistical analyses that determine a bound on the population rate $\dot{n}_S$ vs. the source strength h_0 based on these are described in section 5.

2.2 Population model to source events

As we proceed from observation to interpretation we will need to calculate intermediate quantities that characterize the radiation from our source and population model. In order to follow how those intermediate quantities depend on our source and population model it is useful to give names to different elements of the model names. Let

$$\begin{aligned}
\mathcal{I} &= \left(\begin{array}{l}\text{A source population model, exclusive of the}\\ \text{source rate } \dot{n}_S \text{ and intrinsic luminosity } h_0\end{array}\right) && (1)\\
\dot{n}_S &= \left(\begin{array}{l}\text{rate of gravitational wave events from}\\ \text{the population characterized by } \mathcal{I}\end{array}\right) && (2)\\
h_0 &= \left(\begin{array}{l}\text{intrinsic strain amplitude associated}\\ \text{with sources in the population model } \mathcal{I}\end{array}\right). && (3)
\end{aligned}$$

Here and below we assume that $\mathcal{I}$ is fixed. Our eventual goal is to bound $(\dot{n}_S, h_0)$.

Gravitational wave bursts incident on a gravitational wave detector array can be characterized by their propagation direction and waveform in each of two polarization modes. Since here we assume a single component source model we can, in the context of our model, fully characterize the wave burst incident on the detector in terms of its direction of propagation and the wave amplitude in each of the two polarization modes. Write the full set of parameters that characterize a gravitational wave burst incident on the detectors as

$$\vec{h} = \left(\begin{array}{l}\text{Parameters describing gravitational}\\ \text{waves incident on detector array}\end{array}\right). \tag{4}$$

(In a more complicated source model additional parameters may be needed to fully characterize a wave burst incident on the detector array and $\vec{h}$ would represent this fuller set.)

The population $\mathcal{I}$ leads to a distribution of events at the detector that we write as $p(\vec{h}|h_0, \mathcal{I})$:

$$p(\vec{h}|h_0, \mathcal{I}) = \left(\begin{array}{l}\text{probability of source event characterized}\\ \text{by } \vec{h} \text{ given population } \mathcal{I} \text{ and } h_0\end{array}\right) \tag{5}$$

Note how we have explicitly kept track of the dependency of this distribution on both h_0 and $\mathcal{I}$).

2.3 From source event to foreground event

Not every source event leads to foreground event. Whether it does or does not depends on the analysis method, including thresholds, etc., and the detector noise character and calibration. Define

$$\begin{aligned}
\mathcal{J} &= \left(\text{ Analysis pipeline that identifies and characterizes events }\right), && (6)\\
\mathcal{K} &= \left(\text{ Detector calibration and noise character }\right). && (7)
\end{aligned}$$

Now define the probability that a source event $\vec{h}$ leads to an to an observed event:

$$\epsilon(\vec{h}|\mathcal{I}\mathcal{J}\mathcal{K}) = \left(\begin{array}{l}\text{probability that the source event } \vec{h} \text{ gives}\\ \text{rise to a foreground event in the detector}\\ \text{characterized by } \mathcal{K} \text{ and processing pipeline}\\ \text{characterized by } \mathcal{J}\end{array}\right) \tag{8}$$

Note that $\epsilon(\vec{h}|\mathcal{IJK})$ depends on the detector noise $\mathcal{K}$, processing pipeline $\mathcal{J}$, and source population model $\mathcal{I}$ (which gives meaning to the amplitudes in the polarization modes), though not the rate or the intrinsic source amplitude h_0.[1] With these specified, however, $\epsilon(\vec{h}|\mathcal{IJK})$ is readily determined by simulation.

Also important is the total detection efficiency: the fraction of source events that lead to observed events. The total efficiency is readily calculated from the distribution of source events $p(\vec{h}|h_0,\mathcal{I})$ and the event detection efficiency $\epsilon(\vec{h}|\mathcal{IJK})$:

$$\epsilon(h_0\mathcal{IJK}) = \int \epsilon(\vec{h}|\mathcal{IJK})p(\vec{h}|h_0,\mathcal{I})\, d^n h, \tag{9}$$

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. From the total detection efficiency we can calculate the foreground event rate $\dot{n}_F$ in terms of the source event rate $\dot{n}_S$:

$$\begin{aligned} \dot{n}_F &= \left(\text{ rate of observed foreground events }\right) && (10)\\ &= \dot{n}_S \epsilon(h_0\mathcal{IJK}) && (11) \end{aligned}$$

Each candidate event is characterized by some set of parameters that we denote $\vec{H}$:

$$\vec{H} = \begin{pmatrix} \text{Parameters describing gravitational wave} \\ \text{event identified in the detector array} \end{pmatrix}. \tag{12}$$

Note that we refer to a detector array. We assume that, having different detectors, we use coincidence with time-of-arrival and amplitude constraints to characterize events. At the very least $\vec{H}$ will include some measure of the event amplitude; additionally, through time of arrival measurements it may include also a set of possible event origins on the sky. With enough detectors it may be that $\vec{H}$ includes the amplitudes in both polarizations, a single location on the sky, and other information about the character of the burst. The dimensionality and details of the parameterization $\vec{H}$ depends on the character of the detector array (e.g., number of interferometers, their positions and orientations) and the nature of the analysis that identifies an event. For LIGO and TAMA we will need to be concerned about consistency between the burst amplitudes in the different detectors and the burst "arrival time" in the different detectors.

A critical relationship for us is the one between actual events, described by $\vec{h}$ and detected events, described by $\vec{H}$. Let $q(\vec{H}|\vec{h},\mathcal{IJK})$ be the probability that the foreground event associated with source event $\vec{h}$ is characterized by $\vec{H}$:

$$q(\vec{H}|\vec{h}\mathcal{IJK}) = \begin{pmatrix} \text{probability that observed event associated} \\ \text{with actual event } \vec{h} \text{ is characterized by } \vec{H} \end{pmatrix}. \tag{13}$$

In the absence of noise we would expect that there is a one-to-one mapping from $\vec{h}$ to $\vec{H}$. Noise makes this a one-to-many mapping. The probability $q(\vec{H}|\vec{h}\mathcal{IJK})$ can be thought of as the uncertainty in the determination of the properties $\vec{H}$ associated with the event $\vec{h}$. Note that $q(\vec{H}|\vec{h}\mathcal{IJK})$

[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 6.4.

depends on the source and source population model ($\mathcal{I}$), the nature of the analysis ($\mathcal{J}$), and the nature of the detector noise and calibration ($\mathcal{K}$). Once these are specified, however, $q(\vec{H}|\vec{h}\mathcal{I}\mathcal{J}\mathcal{K})$ is readily determined, along with $\epsilon(\vec{h},\mathcal{I}\mathcal{J}\mathcal{K})$, by simulation.

The distribution in $\vec{H}$ of the foreground contribution to the candidate event distribution is thus

$$P_F(\vec{H}|h_0\mathcal{I}\mathcal{J}\mathcal{K}) = \left(\begin{array}{l}\text{probability of observing the event described by}\\ \vec{H}\text{ given the population characterized by }\mathcal{I}, h_0\end{array}\right) \tag{14}$$

$$\propto \int d^n h\, q(\vec{H}|\vec{h}\mathcal{I}\mathcal{J}\mathcal{K})\epsilon(\vec{h}|\mathcal{I}\mathcal{J}\mathcal{K})p(\vec{h}|h_0,\mathcal{I}) \tag{15}$$

Note that $P_F(\vec{H}|h_0,\mathcal{I}\mathcal{J}\mathcal{K})$ is completely described by the population model through $p(\vec{h}|h_0,\mathcal{I})$ and the analysis pipeline $\mathcal{J}$ as characterized by simulation.

3 Background distribution and event rate

Candidate events may arise from the source population, in which case they are drawn from the distribution $P_F(\vec{H}|h_0,\mathcal{I}\mathcal{J}\mathcal{K})$, or from environmental or instrumental artifacts. We refer to the distribution of events associated with environmental or instrumental artifacts as the background distribution

$$P_B(\vec{H}|\mathcal{J}\mathcal{K})d^nH = \left(\begin{array}{l}\text{fraction of background}\\ \text{events in the }n\text{-ball }d^nH\end{array}\right). \tag{16}$$

Note that the background distribution depends on the detector noise and calibration ($\mathcal{K}$) and the method used to identify events ($\mathcal{J}$).

The background distribution event distribution can be determined using time-delay analysis, under the assumption that background events in each individual detector are uncorrelated at zero-delay. In addition to determining the background distribution, time-delay analysis also determines the expected rate of background events:

$$\dot{n}_B = \left(\begin{array}{l}\text{expected rate of background}\\ \text{events of any amplitude}\end{array}\right). \tag{17}$$

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" h_0 and a source event rate $\dot{n}_S$, and the observation is a set of N observed events $\vec{H}$:

$$\mathcal{H} = \left\{\vec{H}_n : n = 1\ldots N\right\}. \tag{18}$$

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 $\epsilon(h_0,\mathcal{I}\mathcal{J}\mathcal{K})$ (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, $\alpha \in [0,1)$:

$$\epsilon(h_0\mathcal{I}\mathcal{J}\mathcal{K})\dot{n}_S = \dot{n}_F = \dot{n}_B\frac{\alpha}{1-\alpha}. \tag{19}$$

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

$$P(\vec{H}|h_0\alpha\mathcal{I}\mathcal{J}\mathcal{K}) = (1-\alpha)\, P_B(\vec{H}|\mathcal{J}\mathcal{K}) + \alpha P_F(\vec{H}|h_0\mathcal{I}\mathcal{J}\mathcal{K}). \tag{20}$$

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

$$P\left(\mathcal{H}|h_0T\dot{n}_B\alpha\mathcal{I}\mathcal{J}\mathcal{K}\right) = P(N|\mu)\begin{cases} 1 & N=0 \\ \prod_{n=1}^{N} P(\vec{H}_n|h_0\alpha\mathcal{I}\mathcal{J}\mathcal{K}) & N>1 \end{cases} \tag{21}$$

where

$$P(N|\mu) = \frac{\mu^N}{N!}e^{-\mu} \tag{22}$$

is the Poisson distribution and

$$\mu = T\left[\dot{n}_B + \dot{n}_S\epsilon(h_0\mathcal{I}\mathcal{J}\mathcal{K})\right] = \frac{T\dot{n}_B}{1-\alpha} \tag{23}$$

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

In the next section we describe both Bayesian and Frenquentist analyses that use equations 21 and 20 to bound $(h_0, \dot{n}_S)$.

5 Rate vs. Strength

5.1 A Bayesian analysis

For our Bayesian analysis we begin with the likelihood as given in equation 21. Introduce a prior for $(h_0, \dot{n}_S)$, which determines (through the known $\dot{n}_B$ and $\epsilon(h_0\mathcal{I}\mathcal{J}\mathcal{K})$) the prior on α:

$$\begin{aligned} P(h_0,\dot{n}_S) &= \begin{pmatrix} \text{a priori probability density expressing degree} \\ \text{of belief that } h_0,\, \dot{n}_S \text{ take on particular values} \end{pmatrix} && (24) \\ P(h_0,\alpha) &= \begin{pmatrix} \text{a priori probability density expressing degree} \\ \text{of belief that } h_0,\, \alpha \text{ take on particular values} \end{pmatrix} && (25) \\ &= P(h_0,\dot{n}_S)\frac{\dot{n}_B}{\epsilon(h_0\mathcal{I}\mathcal{J}\mathcal{K})(1-\alpha)^2}. && (26) \end{aligned}$$

The observation changes our belief that h_0 and α (and, thus $\dot{n}_S$) take on particular values. This a posteriori probability density is, through Bayes law,

$$P(h_0,\alpha|T\dot{n}_B\mathcal{H}\mathcal{I}\mathcal{J}\mathcal{K}) \propto P(h_0,\alpha)P(\mathcal{H}|h_0T\dot{n}_B\alpha\mathcal{I}\mathcal{J}\mathcal{K}). \tag{27}$$

The proportionality constant, which is formally $1/P(\mathcal{H}|T\dot{n}_B\mathcal{I}\mathcal{J}\mathcal{K})$, is obtained by normalizing the probability density $P(h_0,\alpha|T\dot{n}_B\mathcal{H}\mathcal{I}\mathcal{J}\mathcal{K})$.

Determination of the posterior depends on the choice of prior $P(h_0, \dot{n}_S)$. The choice of prior is important for setting upper limits and, correspondingly, for marginal detection. When a population is clearly observable in the data it will, through the strong peakedness it leads to in the likelihood, overcome any prejudices we impose in any reasonable prior.

Nevertheless, at present we are in the regime of upper limit setting and/or marginal detection of a gravitational wave source population and our choice of prior is, consequently, meaningful. Still, it is not so consequential that the validity of a result turns on the precise choice. In the case of an upper limit, the choice of prior plays a role similar to the choice of ranking principle in a Frequentist analysis leading to a confidence interval or upper limit. Most important is that in the presentation of the results of an analysis we should

- state, up front, our methodology, which for Bayesian includes the choice of prior and for the Frequentist the choice of ranking principle;
- report the likelihood, which permits the audience to make their own choices and draw their own conclusions;
- be clear when the observation is, in fact, uninformative, rather than simply draw the confidence interval, upper limit, or credible set and leave it at that.

5.2 A Frequentist analysis

Having determined the likelihood $P(\vec{H}|h_0\alpha\mathcal{I}\mathcal{J}\mathcal{K})$, the probability of observing the single event $\vec{H}$, we can proceed to find from $\mathcal{H}$ the *joint* bound on $(h_0, \dot{n}_S)$: i.e., instead of finding the bound on the rate assuming the source distribution characterized by h_0 we can find the region in $(h_0, \dot{n}_S)$ space that best explains the observations.

Begin by introducing a partition of the parameter space spanned by $\vec{H}$. For example,

- if $\vec{H}$ is just the event amplitude, then introduce a partition in amplitude;
- if $\vec{H}$ is the amplitude in two different polarizations and the wave propagation direction, then introducing a partition in the amplitude in each polarization and a partition on the sphere for the wave propagation direction.

Bin the individual events in the observation $\mathcal{H}$ according to this partition of the parameter space.

From our knowledge of $P(\vec{H}|h_0T\dot{n}_B\alpha\mathcal{I}\mathcal{J}\mathcal{K})$ and the observation duration T, and under the assumption that the events are independent, we know the expected number of events in each bin introduced above. Form $\chi^2(\mathcal{H}|h_0T\dot{n}_B\alpha\mathcal{I}\mathcal{J}\mathcal{K})$, the χ^2 statistic for the observation $\mathcal{H}$ as binned. Note that, because we expect a particular number of events (we have specified T, $\dot{n}_B$ and $\dot{n}_S$) the χ^2 statistic is drawn from the χ^2 distribution with N degrees of freedom, where N is the number of single events in $\mathcal{H}$.

We can now ask what hypotheses $(h_0, \dot{n}_S)$ lead to $\chi^2(\mathcal{H}|h_0T\alpha\mathcal{I}\mathcal{J}\mathcal{K})$ such that the probability of obtaining this χ^2 is greater than, e.g., 90%. While not a confidence interval in the usual sense, it has a similar interpretation as the range of hypotheses for which the observation is likely in the χ^2 sense.

This approach has the added advantage that it provides, coincidentally, a measure of goodness-of-fit: in particular, if there is no set of hypotheses that include the observation as likely, no "confidence region" will be returned.

6 Discussion

6.1 Introduction

In the first sub-section below we highlight several issues that need to be addressed to realize the analysis describe herein. In ensuing sub-sections we address how we overcome issues of noise and/or calibration non-stationarity, etc.

6.2 Realizing the analysis

6.2.1 Source and population model

The analysis described here interprets gravitational events in terms of a canonical source and distribution of sources in space. The simplest astrophysically motivated model might assume that the source is axisymmetric, in which case there is a single wave-shape and the radiation in each polarization mode depends on the wave shape and the angle between the symmetry axis and the wave propagation direction in a well-defined manner.

6.2.2 Coincidence

The analysis described here begins with events that have been identified as coincident in the LIGO and TAMA detectors. The better job we do with the coincidence cut the more sensitive our analysis will be. With four detectors at three sites we can, in principle, make use of time of flight and event amplitude information to make our coincidence quite strong against background. In particular:

- Events at H1 and H2 should have calibrated amplitudes that are in the ratio 2:1;
- A gravitational wave event propagating in a given direction will, when incident on our detector array, lead to arrival times and calibrated wave strengths that have a particular relation amongst each other. The wave strength relationship depends, in varying degrees, on the source model: in particular, there is no room for variation between H1 and H2, some little room for variations between H1 and L1, and considerable (but still limited) room for variation between any of the LIGO detectors and TAMA. We can insist that coincident events share amplitudes and times-of-arrival that are consistent with a single wave propagation direction.
- Owing to the large variation in sensitivity among the different detectors, sources that would not be observable in one detector may be observable in others. If we allow for this possibility, we can set thresholds for each detector based on their intrinsic sensitivity and have, in the end, a more sensitive search.
- We will need to develop a method that conflates the separate amplitude measurements at the different detectors into a single amplitude measurement that characterizes the coincident events.

6.2.3 Simulations

Simulations will play an especially important role in enabling the joint LIGO/TAMA analysis described here. In particular,

- Propagating the source model through the coincidence step described above will require simulations that add identical signals drawn from the source population to the output of each detector.
- Determining the time-of-arrival difference windows that are important in the coincidence step will require simulations;
- Determining the amplitude ratio windows that are important in the coincidence step will require simulations.

In order to be effective these simulations will require close and careful agreement on the source and population model.

6.2.4 Background determination

The analysis described here depends on an estimate of the background event rate $\dot{n}_B$ and distribution $P_B(\vec{H})$. Under the assumption that events that contribute to the background arise coincidence among uncorrelated events at each site, time-delays between sites (i.e., *not* between H1 and H2) can be used to estimate these quantities.

6.3 A "weak" analysis

Having determined an expected background distribution, we can immediately consider a weaker, but simpler, analysis than has been described above. In particular we can ask whether the observed events are consistent with the expected background event number and distribution. For a Bayesian analysis, the likelihood that the observed distribution is background is given by equation 21 with α equal to 0 (i.e., ignoring the source model entirely); for a Frenquentist analysis, the distribution and number of observed events can be compared, using a χ^2 test, with the expected distribution and number from the background. A test of this kind does not place a limit on the rate or strength of a gravitational wave source population; rather, it makes the simple statement that the observed distribution and number are, or are not, consistent with the expected background.

6.4 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 $\mathcal{H}$ 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 $\mathcal{H}$ into M disjoint sub-observations $\mathcal{H}_k$, with the union of the $\mathcal{H}_k$ equal to $\mathcal{H}$, such that all the events in $\mathcal{H}_k$ 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: $P\left(\mathcal{H}_k|h_0, t_k, \dot{n}_{B,k}, \dot{n}_S\right)$. *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

$$P\left(\mathcal{H}|h_0, \{t_k, \dot{n}_{B,k}\}, \dot{n}_S\right) = \prod_{k=1}^{M} P\left(\mathcal{H}_k|h_0, t_k, \dot{n}_{B,k}, \dot{n}_S\right). \tag{28}$$

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.

6.5 Background rate uncertainty

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

$$P_B\left(\dot{n}_B\right) d\dot{n}_B = \left(\text{ degree of belief that } \dot{n}_B \text{ is in } [\dot{n}_B, \dot{n}_B + d\dot{n}_B).\ \right) \tag{29}$$

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

$$P(\mathcal{H}|h_0, T, \dot{n}_S, \mathcal{I}\mathcal{J}) = \int d\dot{n}_B P_B(\dot{n}_B) P(\mathcal{H}|h_0, \dot{n}_B, \dot{n}_S, \mathcal{I}\mathcal{J}) \tag{30}$$

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.