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

**Document Type** **LIGO-T030181-00-Z** 2003/08/28

# A Veto Selection Criteria for Gravitational Wave Burst Searches

Lee Samuel Finn

*Distribution of this draft:*

LIGO Scientific 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 2003/08/28

**Abstract**

Veto selection for burst analysis has focused on an ad hoc comparison of the number of vetoed events to the loss of livetime associated with the vetos. Here we propose a different approach, which focuses on the effectiveness of the veto and involves only a dimensionless ratio that is greater than unity when the veto is effective.

$Id: T030181.tex,v 1.3 2003/08/28 13:41:25 lsf Exp $

# Contents

**1 Introduction** **2**

**2 Nomenclature** **3**

**3 An ineffective veto** **3**

**4 An effective veto** **4**

**5 When is $\mathcal{R}$ significantly greater than unity?** **4**

**6 Application: Is a veto effective?** **5**

**7 Tuning a veto** **6**

**8 Conclusion** **6**

**A A MATLAB program for evaluating the $\mathcal{R}$ threshold and distribution** **6**

# 1 Introduction

Veto selection for burst analysis has focused on an ad hoc comparison of the number of vetoed events to the loss of livetime associated with the vetos. While the desiderata of "more vetoed events than vetoed livetime" is compelling, making it quantitative as stated involves the comparison of a dimensioned quantity (vetoed livetime) with a differently dimensioned (actually, undimensioned) quantity (number of vetoed events).

Here we propose a different approach, which involves only a dimensionless ratio that is greater than unity when the veto is effective. A veto divides the events and total livetime into two categories: vetoed events and vetoed livetime (sometimes referred to as deadtime) and unvetoed events and unvetoed livetime. If the number of vetoed events in the vetoed livetime is greater than the number of unvetoed events in the unvetoed livetime the veto is effective. In the rest of this note we develop this idea into a quantitative method for tuning a prospective veto for greatest effectiveness and judging whether the tuned veto is, in fact, effective.

# 2 Nomenclature

Before the application of the veto there are events and a livetime:

$$\begin{aligned} N &= (\text{Number of events before veto application}) && (1) \\ T &= (\text{Livetime before veto application}) && (2) \end{aligned}$$

A veto divides the events and the livetime into two categories: those events and that livetime vetoed, and those events and that livetime not vetoed:

$$\begin{aligned} N_0 &= (\text{Number of events unvetoed}) && (3) \\ T_0 &= (\text{Livetime unvetoed}) && (4) \\ N_V &= (\text{Number of events vetoed}) && (5) \\ T_V &= (\text{Livetime vetoed}) && (6) \end{aligned}$$

Events (and livetime) are either vetoed or unvetoed:

$$\begin{aligned} N &= N_0 + N_V && (7) \\ T &= T_0 + T_V && (8) \end{aligned}$$

Finally define the rate of events before vetoing, in the vetoed livetime, and in the livetime remaining after vetoing:

$$\begin{aligned} R &\equiv \frac{\overline{N}}{\overline{T}} && (9a) \\ R_V &\equiv \frac{\overline{N}_V}{\overline{T}_V} && (9b) \\ R_0 &\equiv \frac{\overline{N}_0}{\overline{T}_0}. && (9c) \end{aligned}$$

# 3 An ineffective veto

Consider a completely ineffective veto: i.e., a veto that is completely uncorrelated with events. In an average over an ensemble of experiments the number of events vetoed, $\overline{N}_V$, will be in proportion to the livetime vetoed, $\overline{T}_V$: i.e.,

$$\overline{N}_V = \overline{N}\frac{\overline{T}_V}{T}. \tag{10}$$

(Since vetos are assumed uncorrelated with events $\overline{NT_V}$ is equal to $\overline{T}\,\overline{N}_V$.) The ratio of the rate of events in the vetoed time $T_V$ to the rate of events in the remaining time are equal:

$$\begin{aligned} R_0 &= \frac{\overline{N}_0}{\overline{T}_0} && \text{(11a)} \\ &= \frac{\overline{N}-\overline{N}_V}{T-\overline{T}_V} && \text{(11b)} \\ &= \frac{\overline{N}\left(1-\overline{T}_V/T\right)}{T-\overline{T}_V} && \text{(11c)} \\ &= \frac{\overline{N}}{T} && \text{(11d)} \\ &= \frac{\overline{N}_V}{\overline{T}_V} && \text{(11e)} \\ R_0 &= R_V && \text{(11f)} \end{aligned}$$

# 4 An effective veto

From our discussion of an ineffective veto it should be clear, in contrast, that an effective veto should remove more events than livetime *in comparison with the remainder*: i.e., a figure of merit for vetos is

$$\mathcal{R} = \frac{N_V}{T_V}\frac{T_0}{N_0}. \tag{12}$$

For an effective veto $\mathcal{R}$ should be significantly greater than unity and the greater $\mathcal{R}$ the greater the veto effectiveness.

Note that the figure of merit $\mathcal{R}$ is independent of the actual number or rate of events, or the number or rate of events vetoed: a veto is judged effective if it sets-aside more events than would be expected by chance.

# 5 When is $\mathcal{R}$ significantly greater than unity?

An effective veto has a mean value of $\mathcal{R}$ significantly greater than unity. In any particular instance of events and vetoes $\mathcal{R}$ will fluctuate from its mean value. To identify when a veto is effective we must determine when $\mathcal{R}-1$ is larger than would be expected owing to statistical fluctuations associated with an ineffective veto. For this purpose we need to determine the probability of observing a given figure of merit $\mathcal{R}$ in the case of an ineffective veto. Knowing that distribution we can ask for that $\mathcal{R}_p$ such that $\mathcal{R} > \mathcal{R}_p$ with probability $p$. Choosing $p$ appropriately small gives us a threshold that our prospective veto must cross before we judge it effective.

In an observation of duration $T$ the distribution of fluctuations in $\mathcal{R}$ is governed by

$$\begin{aligned} \overline{N}_0 &= \text{(The expected number of events not vetoed)} && \text{(13a)} \\ \overline{N}_V &= \text{(The expected number of events vetoed)} && \text{(13b)} \\ \overline{T}_V &= \text{(The expected amount of time vetoed)} && \text{(13c)} \end{aligned}$$

A simplifying assumption, which is likely useful for the purpose of identifying effective vetoes, is to take the number of events $N$ to be Poisson distributed and the number of vetoed events to be Binomially distributed:

$$P(N|\overline{N}) = \frac{\overline{N}^N}{N!} e^{-N} \tag{14a}$$

$$P(N_V|q, N) = \frac{N!}{N_V!(N - N_V)!} q^{N_V} (1 - q)^{(N-N_V)}. \tag{14b}$$

where will take

$$\overline{N} = N_V + N_0, \tag{14c}$$

$$q = \frac{T_V}{T}. \tag{14d}$$

This approximation captures the fluctuations associated with the total number of events and the fraction of those incidentally vetoed by the ineffective veto. In this approximation the probability distribution for $\mathcal{R}$ depends on the total number of observed events $N$ and the fraction of the total time vetoed $q$. With these determined the distribution of $\mathcal{R}$ is readily found via a Monte Carlo simulation and the threshold $\mathcal{R}_p$ for any $p$ determined. The appendix shows a MATLAB program that will determine the distribution of $\mathcal{R}$ and the threshold $\mathcal{R}_p$ for given $p$, $q$, and $N$.

# 6 Application: Is a veto effective?

Suppose that we observe $N$ events in our playground data set, which has duration $T$. Upon application of a prospective veto we find a total of $N_V$ events are vetoed at a cost of time $T_V$ lost to the veto.

$$\mathcal{R} = \frac{N_V}{T_V} \frac{T - T_V}{N - N_V}. \tag{15}$$

We have decided that an effective veto is one whose figure of merit is so great that we would not expect it for an ineffective veto more than 5 in 100 similar experiments. Using the MATLAB program in the appendix we determine the threshold $\mathcal{R}_p$ to be given by

$$\mathcal{R}_p = \texttt{rdist}(N, T_V/T, 5/100); \tag{16}$$

If $\mathcal{R} > \mathcal{R}_p$ we judge the veto to be effective.

For example, suppose that

$$N = 13\,502 \tag{17a}$$

$$T = 102.8\,\mathrm{h} \tag{17b}$$

$$N_V = 276 \tag{17c}$$

$$T_V = 1.93\,\mathrm{h}. \tag{17d}$$

Then

$$q = \frac{T_V}{T} = 0.0188 \tag{17e}$$

$$\mathcal{R} = 1.086 \tag{17f}$$

Now suppose we insist that a veto is effective only if the chance that an ineffective veto would give us a value of $\mathcal{R}$ this great or greater is less than 5%. Using the MATLAB function `rdist` from the appendix we find that

$$\texttt{rp} = \texttt{rdist(N,q,0.05)} \tag{17g}$$

The threshold `rp` so obtained is 1.1252, which is greater than $\mathcal{R}$. Correspondingly, we conclude that this veto is *no different than a chance selection* with a p-value of 5%.

# 7 Tuning a veto

Prospective vetoes generally involve one or more parameters that change $T_V$ and $N_V$: e.g., a threshold. We can use the figure of merit described here to simultaneously tune and evaluate a veto for greatest effectiveness.

Let $\theta$ denote the parameters, which control the veto, that we wish to tune. The figure of merit on any given data set is then a function of $\theta$: i.e.,

$$\mathcal{R} = \mathcal{R}(\theta). \tag{18}$$

Randomly the data set available for evaluating the veto into two sets, $\mathcal{T}$ (for train) and $\mathcal{V}$ (for validate). On set $\mathcal{T}$ chose the $\theta$ that maximizes $\mathcal{R}(\theta)$: i.e.,

$$\theta_{\max} = \theta \quad \text{such that} \quad \mathcal{R}(\theta_{\max}) \geq \mathcal{R}(\theta) \forall \theta. \tag{19}$$

Having selected $\theta_{\max}$ in this way, evaluate $\mathcal{R}(\theta_{\max})$ on the validation set $\mathcal{V}$ using the protocol described in section 6 above. The veto with parameters $\theta_{\max}$ is effective if it is so judged by this protocol; otherwise it is not effective and is unlikely to be effective for any other value of $\theta$.

# 8 Conclusion

We have defined a quantitative formulation of "effectiveness" for vetoes based on the criterion that an effective veto should identify events with greater frequency than then occur in the unvetoed data. The corresponding figure of merit is independent of the number of events or the event rate. Better vetoes, in the sense of greatest rate of events vetoed compared to the rate of events remaining, have uniformly higher figure of merit then less effective vetoes, permitting vetoes to be compared against each other or tuned. We offer an ansatz for estimating a threshold on the figure of merit such that vetoes whose figure of merit exceeds the threshold actually have a positive effect, compared to random chance and illustrate with an example. Finally, we describe how to tune a veto for greatest effectiveness in the sense of greatest rate of events vetoed compared to the rate of events remaining.

# A A MATLAB program for evaluating the $\mathcal{R}$ threshold and distribution

For a discussion of this program and its use see sections 5 and 6.

```
function varargout = rdist(n,q,p)
% RDIST - evaluate the threshold and distribution for the veto
% figure of merit descried in T030181
%
% rp = rdist(n,q,p)
% [rp,dr] = rdist(n,q,p)
%
% n      total number of events, vetoed and unvetoed
% q      fraction of total observation time vetoed
% p      threshold probability
%
% rp     threshold value of R such that R > rp with probability p
% dr     cumulative distribution R > dr(k,2) with probability dr(k,1)
%
% Requires statistics toolbox

% $Id: T030181.tex,v 1.3 2003/08/28 13:41:25 lsf Exp $

% basic error checking
if (nargout < 1)
    error('usage: [rp,dr] = rdist(n,p,q) or rp = rdist(n,p,q)');
end
if (n ~= round(n) | n < 1)
    error('integer n must be greater than 0');
end
if (~(q < 1 & q > 0))
    error('probabilty q must be in (0,1)');
end
if (~(p < 1 & p > 0))
    error('probabiltiy p must be in (0,1)');
end

% To get a tail that resolves probabilty to better than 10%
% need at least 100/p draws
nsim = ceil(100/p);

% draw total number of events
ntot = poissrnd(n,nsim,1);
% draw number of vetoed events
nvet = binornd(ntot,q,nsim,1);
% evaluate R
r = nvet./((ntot - nvet)*q);

% find threshold
varargout{1} = prctile(r,100*(1-p));
```

```
if (nargout > 1)
    % make cumulative distribution
    r = sort(r);
    dr = zeros([2,length(r)]);
    dr(1,:) = (1:length(r))/length(r);
    dr(2,:) = r(:)';
    varargout{2} = dr;
end

return
```