INSTITUTE OF PHYSICS PUBLISHING CLASSICAL AND QUANTUM GRAVITY

Class. Quantum Grav. **20** (2003) S633–S643 PII: S0264-9381(03)62479-2

Status of the search for gravitational wave bursts with the LIGO detectors

L Cadonati and E Katsavounidis for the LIGO Scientific Collaboration's Bursts Working Group

Massachusetts Institute of Technology, LIGO Laboratory, NW17-161 Cambridge, MA 02139, USA

E-mail: cadonati@ligo.mit.edu and kats@ligo.mit.edu

Received 23 March 2003
Published 7 August 2003
Online at stacks.iop.org/CQG/20/S633

Abstract

The LIGO Scientific Collaboration has been preparing for the search for burst-like events in the first LIGO science data. We present an overview of the goals of the burst search and provide a description of the different methods used to generate triggers from the raw time series. We also introduce the infrastructure of the prototype burst analysis pipeline. This brings together triggers from the gravitational wave and diagnostics channels in a statistically defined veto strategy, and then uses a multiple interferometer coincidence mechanism to identify candidate burst events. Finally, we discuss various aspects of the burst analysis, such as the identification of the detector stationarity and strategies in the determination of the background, upper limits and efficiency estimation. These are presented as of the time of the 7th GWDAW workshop that took place in Kyoto, Japan in December 2002.

PACS numbers: 95.55.Ym, 04.80.Nm

1. Introduction

The Laser Interferometer Gravitational wave Observatory (LIGO) [1] is comprised of three long-baseline interferometers situated in Livingston, LA and Hanford, WA that will be able to detect gravitational waves at a sensitivity of $h_{\rm rms} \approx 10^{-21}$ at about 100 Hz. The detectors built by a Caltech–MIT collaboration are completed and their commissioning is progressing towards their design sensitivity. A major milestone in the instruments' commissioning and operation was the first science run (which will be referred to as S1 from this point on). S1 lasted 17 days and collected coincidence data from all three LIGO detectors from 23 August 2002 to 9 September 2002. The detectors were operated in their final optical configuration as power-recycled Michelson interferometers with Fabry–Perot cavities. The sensitivities (see figure 1) of the three instruments during S1 were to within a factor of 100 of their design goals

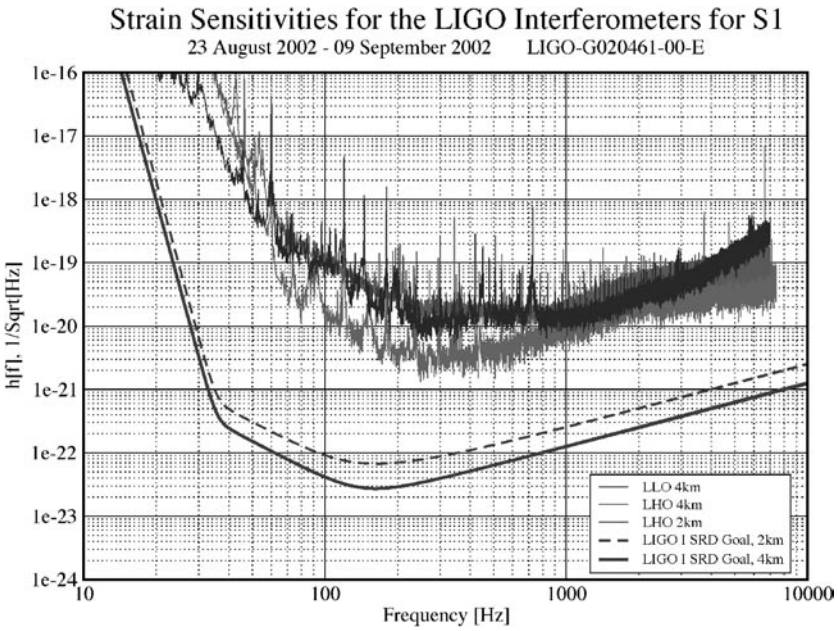

Figure 1. Spectral strain sensitivity of the three LIGO interferometers during the first science run (S1).

with the best one presented by the Livingston-4 km interferometer at $\sim 3 \times 10^{-21}$ $\mathrm{Hz}^{-1/2}$ at 300 Hz. The total observation time of S1 was 408 h. During this time, the three interferometers operated simultaneously for 95.7 h.

The S1 data reflect the most sensitive interferometric data collected so far over a broad band of frequencies taken by detectors in coincidence. Simultaneously with LIGO's S1, the GEO-600 [2] project in Germany collected interferometric data with the detector in power-recycled mode.

The LIGO Scientific Collaboration (LSC) [3] has analysed data from the S1 run, simultaneously interested in searching for any detectable gravitational wave signal and in setting upper limits if none is found, as astrophysical wisdom suggests [4]. This investigation was mostly aimed at establishing the methods and infrastructure to be adopted in the analysis of science data that the LIGO detectors will collect over the next two years.

Within the LSC, the search for gravitational wave bursts is pursued by the Bursts Working Group [5]. Reports and final results from the analysis of the S1 data by the Bursts Working Group are in their final stages of preparation and they will become available soon [6].

2. S1 goals of the burst search

The goal of the Bursts Working Group is to look for short transients (typically less than 1 s) of gravitational radiation without necessarily knowing their waveform and spectrum. This includes burst signals from supernovae and black-hole mergers, for which the physics and computational implications are complex enough that any analytical calculation of the expected waveforms is extremely difficult. While detailed knowledge of a signal waveform would have allowed the use of optimal matched filtering techniques, the Bursts Working Group

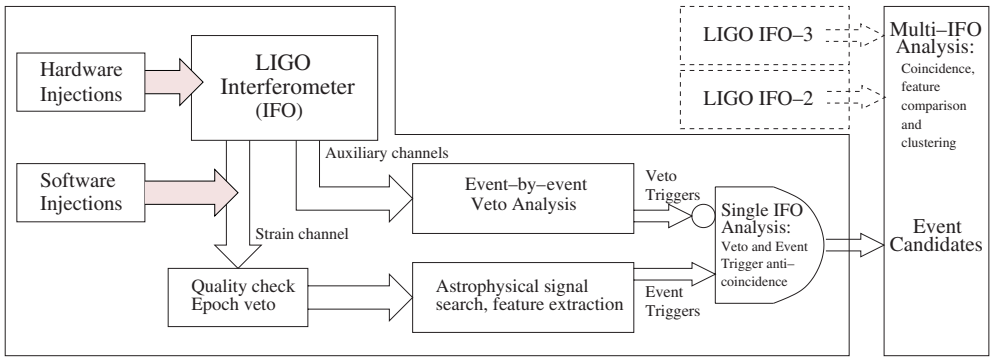

Figure 2. The burst search pipeline. A per-interferometer (IFO) search establishes event triggers in anticoincidence with any possible veto trigger. The multi-IFO analysis combines the information from the individual IFOs and requires their temporal coincidence. Additional feature comparison is performed whenever the search engines yield such information. Monte Carlo techniques based on hardware and software injection are then used to establish the sensitivity of the search to bursts of given strength.

has chosen to pursue a general, model-independent investigation that explores anomalies in the interferometer strain data using general time-domain and time–frequency-domain search techniques. This allows LIGO's quest for bursts to be kept open to any *unanticipated* source of gravitational radiation.

An additional focus of the Bursts Working Group is to look for correlations between gravitational wave bursts and external triggers from neutrino telescopes [7] or γ-ray burst (GRB) detectors. A number of GRB progenitors are plausible gravitational wave burst emitters; a comparison of the correlation function of the LIGO detectors immediately before a GRB (*on source*) and at random times (*off source*) may statistically establish their association [8]. For S1, the Bursts Working Group has focused on the implementation of the analysis technique for such an *externally triggered* search and exercised it on the single GRB event coincident with LIGO's S1 dataset. A full-scale, end-to-end analysis of the sensitivity to GRB–gravitational wave correlations will be pursued in future LIGO runs.

In the following sections, we will focus on the implementation details of the more general, untriggered search.

3. The burst search pipeline

The burst analysis pipeline has been designed to accommodate a generic search for transients using the LIGO detectors and it is depicted in figure 2.

The data stream from the LIGO interferometers (IFOs) comprises more than 5000 channels that are sampled continuously. No hardware or software triggers are created by the detector's acquisition system: all time series are written onto disk and are processed later in software. The time series from the demodulated photodiode at the interferometer antisymmetric port ('AS_Q') records any differential length change in the arms induced by the passage of a gravitational wave. We refer to it as the *gravitational wave channel* or *strain channel* (one such channel per IFO). This time series is processed by the astrophysical burst search algorithms, yielding *event triggers*. Several auxiliary read-back channels of the servo control systems and instruments monitoring the physical environment are simultaneously recorded. These channels are used to veto transients of non-astrophysical origin.

The burst pipeline generates on a per-IFO basis a sequence of events reflecting the anticoincidence of the *event* and *veto triggers*. At a second stage, events from each IFO are brought together for a multi-IFO analysis where the time coincidence of the events is imposed and a comparison of their extracted features—whenever available—is exercised. Hardware and software injections are used to establish the efficiency of the pipeline. Background estimation based on time shifts of individual IFOs event triggers is used to bound the rate of gravitational wave bursts.

A central element in the S1 analysis has been the definition of a *playground* dataset, covering ≈10% of the S1 coincidence data. In order to avoid any statistical biases in the search for bursts, we have restricted to this set any tuning of the search algorithms and veto parameters. Once the analysis was fixed, we applied it to the remaining ≈90% of the S1 coincidence data, exclusive of the playground set.

4. Data stationarity and epoch veto

It is important that the search for transients of gravitational radiation be carried out only on data from sufficiently stationary interferometers. At the most fundamental level, we only analyse data produced when the interferometers are locked in their full configuration and no alignments or other adjustments are taking place. The interferometer operator certifies that each instrument has entered this state (called *science mode*) by pressing a button on the control screen. An automatic system detects when the state has changed either because the system fell out of lock or because an adjustment was made. The intervals of science mode data are recorded in a database.

Several handles on the time-averaged performance of the instruments are provided by the auxiliary channels and by the strain channel. They can be used to establish the quality of the data collected over a period of time and to reject segments during which the instruments were misbehaving. We refer to such data quality cuts as *epoch vetoes*, as they reduce the *live time* of our experiment without any impact on its efficiency. As a first such epoch veto for the burst search, we measure the rms noise in the strain channel in four broad frequency bands, averaged over 360 s periods, and reject lock segments during which the broad-band noise is extraordinarily large.

We define $\kappa_{68\%}$ as the 68 percentile in the distribution of band-limited rms in each interferometer and frequency band and apply a $10\kappa_{68\%}$ cut in the 320–400 Hz band, $3\kappa_{68\%}$ in the 400–600 Hz, 600–1600 Hz and 1600–3000 Hz bands (chosen *ad hoc*). We verified, by simulation, that such signal fluctuations over 6 min periods cannot be induced by short transients but are rather due to non-optimal functioning of the detector. The cut applied in the 320–400 Hz band is less stringent than those in the other bands, since rms fluctuations in this frequency range are only marginally affecting the rate of events from the *trigger generators* (section 5). Once the cuts at different interferometers and in different bandwidths have been combined, the epoch veto affects 33% of the S1 triple-coincidence data. Details of the implementation at the Hanford-4 km are shown in figure 3.

5. Event trigger generators

After being high-passed at 150 Hz and whitened, the time series of the gravitational wave channel is analysed within the LIGO data analysis system (LDAS) [9]. During S1, three search algorithms were employed for the detection of bursts; they are referred to as *event trigger generators* (ETGs).

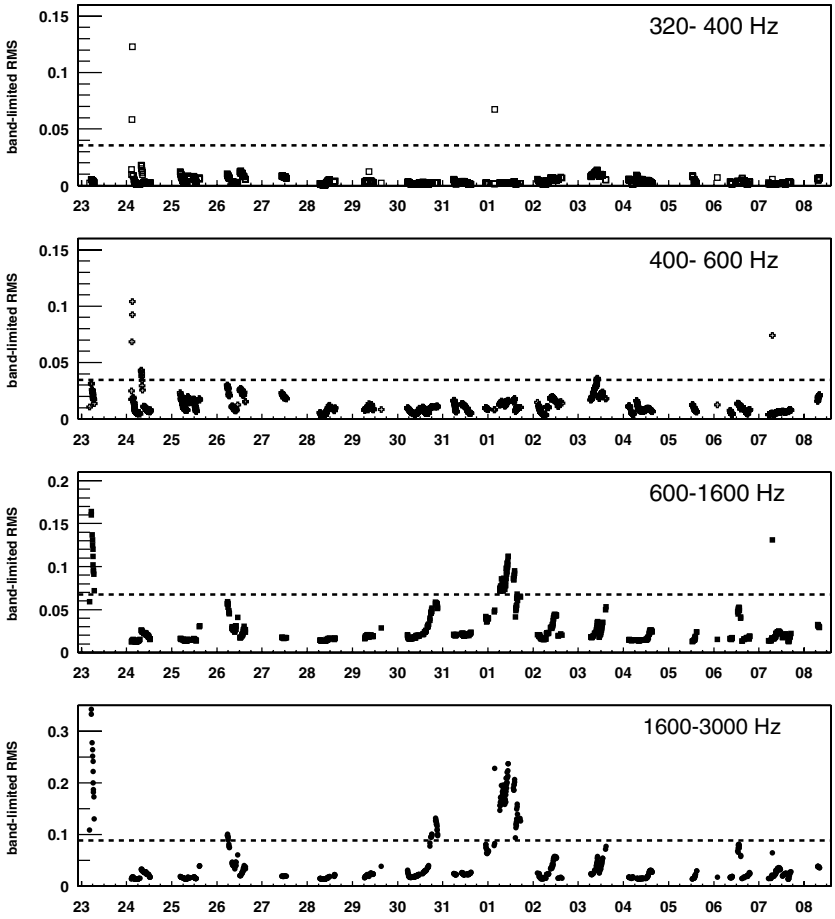

Figure 3. Data stationarity: plots of band-limited rms noise in the gravitational wave channel at the Hanford-4 km, measured over 360 s intervals in the triple-coincidence S1 dataset. On the horizontal axis is the day of the month, from 23 August 2002 to 9 September 2002. We monitored four *ad hoc* bands and rejected lock segments that contain outliers in the band-limited rms distribution. The horizontal dashed lines show where the cut is applied, corresponding to $10\kappa_{68\%}$ in the 320–400 Hz band and $3\kappa_{68\%}$ in the 400–600 Hz, 600–1600 Hz and 1600–3000 Hz bands, where $\kappa_{68\%}$ is the 68 percentile in the rms distribution.

The SLOPE ETG [10] is a time-domain algorithm, described in [11]. It performs a least-squares fit to a line, over a short segment of strain data, and selects candidate events based on the value of the slope. The algorithm reports start time and significance (value of the slope) of the excursion. The TFCLUSTERS ETG is a time–frequency method [13] relying on successive spectrograms, calculated every 0.125 s and thresholded to identify time–frequency pixels (0.125 s $\times$ 8 Hz) with statistically significant excess of power. The algorithm finds clusters of these pixels and reports the total power, central frequency, bandwidth, start time and duration of the cluster. A third method, referred to as POWER [14], uses, like TFCLUSTERS, fast Fourier transforms to calculate the power spectrum for any given start time and duration. The algorithm then compares the power in user-defined time–frequency tiles to the statistical distribution of the noise power. When it finds an excess of power, larger than that expected

from the noise statistical fluctuations, it stores time, duration, frequency, bandwidth, power and confidence of a candidate burst event.

All the triggers identified by each search method, and their features, are stored in the LDAS database, for further pipeline analysis.

6. Use of auxiliary channels and vetoes

Transients in the gravitational wave channel due to instrumental or environmental perturbations are likely to be detected in one of the several auxiliary channels that are recorded along with it: interferometer diagnostics, laser control signals and environmental monitors such as seismometers, microphones, tiltmeters, magnetometers and power line monitors. If one or more of these channels present a particularly strong, statistically significant correlation with the strain channel, we can use their transients to veto a fraction of the noise events detected by the event trigger generators.

Our investigation of auxiliary channels is based on *glitchMon* [15], a transient search algorithm running in the LIGO data monitor tool environment [16]. *GlitchMon* examines a pre-filtered time series for excursions above a fixed threshold. An event is defined by its start time, duration and amplitude. The duration is the time it takes the signal to return and stay below the threshold for at least 0.25 s; the amplitude is the largest instantaneous excursion during the event duration.

We developed a technique to select channels and tune thresholds, based on the statistical correlation between veto triggers and burst candidate events. A successful veto coincides with a significant fraction of the burst candidates at a minimal cost for the live time. For each auxiliary channel, we define the *veto efficiency* as the fraction of burst candidates in coincidence with veto triggers:

$$\varepsilon_V = \frac{N_{\text{vetoed}}}{N_{\text{detected}}}, \tag{1}$$

and the *dead time fraction* introduced by the veto as

$$\tau = \frac{T_D}{T}, \qquad T_D = \sum_i t_i, \tag{2}$$

where T_D is the *dead time*, sum of the individual diagnostic trigger durations.

We use ε_V–τ curves, parametrized by the threshold used in *glitchMon*, to compare the performance of different veto channels and select the most effective. Figure 4 shows examples of such curves for the Hanford-4 km interferometer. The graph compares the performance of two auxiliary channels at vetoing the complete set of playground event triggers or a reduced set of outliers (particularly loud triggers).

The next step consists of tuning the threshold: this is a trade-off between the veto efficiency, its deadtime and the rate of accidental coincidences. For each threshold, we introduce lags between the strain channel and the auxiliary channel time series and measure the number of accidental coincidences obtained as a function of the lag. Figure 5 shows the resulting lag data for the veto channel of choice at the Hanford-4 km. The effect of vetoes on the Hanford-4 km histogram produced by TFCLUSTERS and SLOPE ETGs is shown in figure 6.

To date, only signals from interferometer and laser control signals have shown a good match with any frequent transient phenomena in the strain channel. In the specific case of the S1 analysis, we identified 'AS_DC' (DC current in the photodiode monitoring the antisymmetric port) as the veto channel for the Livingston-4 km interferometer and 'REFL_I' (error signal

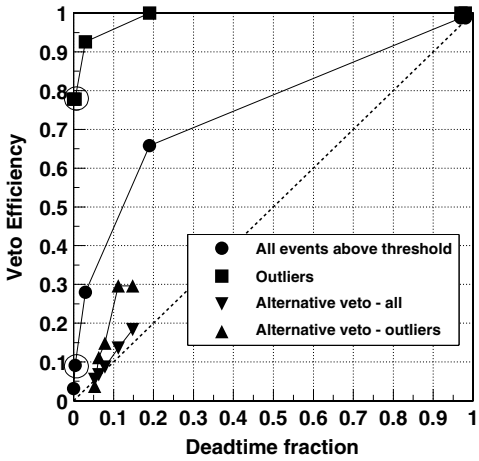

Figure 4. Plots of veto efficiency (ε_V) versus deadtime fraction (τ) at the Hanford-4 km. Each curve is obtained with a different set of veto and TFCLUSTERS event triggers. Each point on a curve corresponds to a different threshold used in the production of the diagnostics triggers. Lowering the threshold traces the curve towards the upper right corner as veto rate and efficiencies increase. An ideal veto lies, in this plot, in the upper left corner: $\varepsilon_V \to 1$ and $\tau \to 0$. The curve for a veto with poor performance is close to the diagonal, a sign that the correlation with the strain channel is purely random (e.g., 10% of the triggers are vetoed at the cost of 10% deadtime). In this example, the filled circles and the downward-pointing triangles have been obtained using all the triggers produced by TFCLUSTERS in the S1 playground dataset. The filled circles are associated with the 'REFL_I' channel (error signal of the common arm length degree of freedom). The triangles are associated with 'REFL_Q' (alternative error signal from the Michelson degree of freedom [17]). The difference in performance becomes more dramatic if we restrict the study to outliers in the TFCLUSTERS power histogram (loud events, chosen by eye, with power $>10^3$ in the top part of figure 6): the squares show the 'REFL_I' curve, the upward-pointing triangles are for 'REFL_Q'. The open circles in the figure indicate the efficiency (9% for the full trigger set, 78% for the outliers) and deadtime fraction (0.4%) for the threshold we ultimately chose, after analysing the lag plots in figure 5.

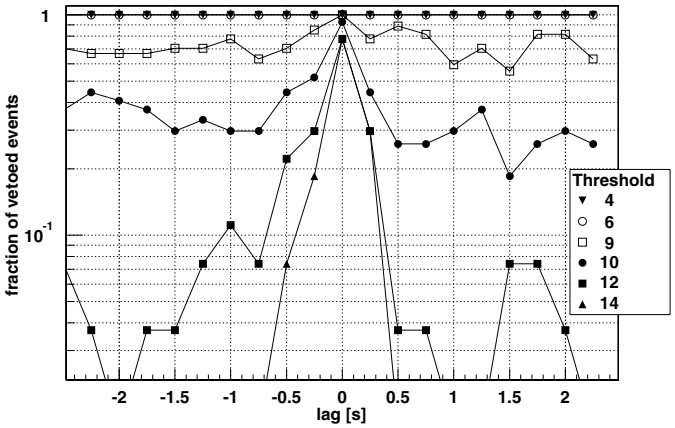

Figure 5. Hanford-4 km veto lag plot, obtained with TFCLUSTERS outliers and triggers on 'REFL_I'. Each set corresponds to a different veto trigger threshold (numerical value reported in the legend). On the horizontal axis is the lag introduced between the strain channel and 'REFL_I'. On the vertical axis is the fraction of vetoed events, which, for lag $= 0$, is the veto efficiency ε_V defined in equation (1). For non-zero lags this quantity is the accidental probability.

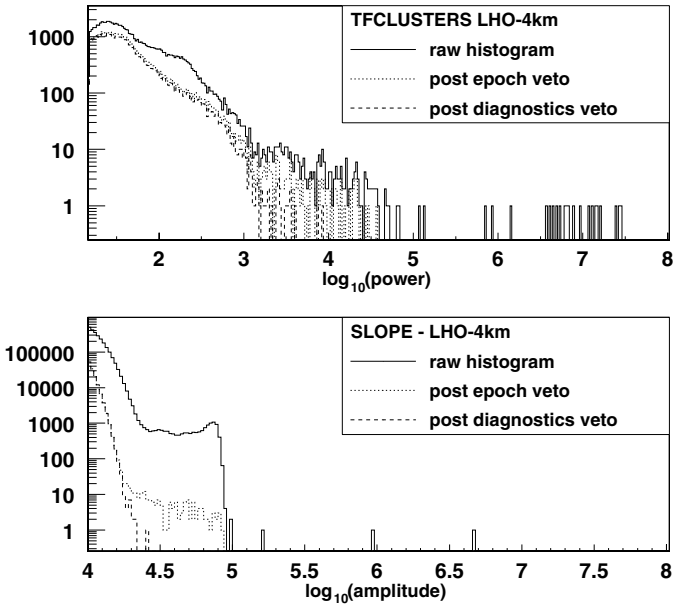

Figure 6. Effect of vetos on the Hanford-4 km event histogram, produced with the TFCLUSTERS (top) and SLOPE (bottom) ETGs. The histogrammed quantity depends on the search algorithm (power for TFCLUSTERS, slope amplitude for SLOPE), and it conveys information on the event strength. We show here the effect of both epoch veto (data selection based on stationary noise, as described in section 4) and the diagnostics veto described in section 6.

of the common arm length degree of freedom) for the Hanford-4 km interferometer, while we found no good veto for the Hanford-2 km.

While it is true that correlations between the gravitational wave channel and the interferometer control signals offer a handle to suppress spurious events, they also make us vulnerable to the possibility that genuine gravitational wave bursts trigger our veto mechanism. For this reason we studied the interferometers' response to induced mirror motions simulating gravitational wave events. We compared the resulting strain channel events to the amplitude of transients induced in the veto channel, with a signal strength ratio test.

We have also measured couplings between the physical environment and the strain channel. None of the environmental channels is an effective veto for the S1 analysis, but as the detectors' sensitivity increases, we expect the coupling to environmental channels to become more and more important, so that in the future we might be able to use these as effective (and safe) vetoes.

7. Coincidence

Additional reduction of the remaining candidate events is obtained by requiring their temporal coincidence in the three LIGO interferometers. A transient of astrophysical origin is expected to yield time-correlated triggers in the three LIGO detectors, subject only to the propagation time between the sites ($\leqslant$10 ms) and any dispersion in the definition of the burst start time by the search algorithms. With the parameters used in the S1 analysis, the time resolution is 125 ms for the TFCLUSTERS algorithm, 3 ms for SLOPE.

We select a list of final candidate events with the requirement that their start time be in temporal coincidence across the three LIGO interferometers. We have determined empirically, by study of simulated signals, that at least 99% of the simulated signals, added simultaneously to the data of each interferometer, would be detected with start times equal to within 500 ms by TFCLUSTERS and 50 ms by SLOPE. This includes an allowance for the light travel time between the sites (up to 10 ms) and the ringing time of the filters applied to the data, which, in the S1 implementation, spreads a 1 ms Gaussian pulse over up to 40 ms.

Additionally, an inter-site correlation of the burst duration, frequency band, amplitude and waveform is to be expected. At present, the only requirement we impose on the final events is a match of the frequency bands covered by the events at the three interferometers. This requirement does not apply to SLOPE, a pure time-domain filter that does not provide frequency information. The exploitation of the full power of the multi-detector burst search is under study and will be implemented in future analysis.

8. Simulations and interpretation

The burst analysis pipeline we have just described yields a measurement of the number of transients detected in coincidence by the three LIGO detectors. This number includes possible gravitational wave bursts (the *foreground*) as well as a number of events attributed to chance coincidence of noise in the three instruments (the *background*). In order to set an upper limit on the foreground, we use the Feldman–Cousins unified approach [18] and compare the measured number of coincident events with the estimated background, with the assumption that both foreground and background are Poisson processes. The background is estimated by introducing several time lags between the time series at different interferometers and averaging the number of measured coincidences. This analysis yields an *uninterpreted* upper limit on the measured rate of bursts.

A full set of Monte Carlo simulations has also been used to inject Gaussian and sine-Gaussian signals of variable strength, width and frequency content onto the actual interferometer time series. For each waveform, we separately evaluate the efficiency for detection through each of the three LIGO detectors and analysis pipelines, as a function of the signal peak amplitude h_0, assuming optimal source orientation and polarization. The efficiencies are essentially 100% for large values of h_0, consistent with noise (and thus 0% efficiency) for small h_0 and transitioning smoothly over an intermediate range of h_0. The results are fitted to a sigmoid curve in $\log_{10}(h_0)$:

$$\varepsilon(h_0) = \frac{1}{1 + \mathrm{e}^{(\log_{10} h_0 - b)/a}}, \tag{3}$$

where $b = \log_{10} h_{1/2}$ determines the peak amplitude at which the efficiency is equal to $1/2$, and a governs the width of the transition from 0 to 1 in $\log_{10}(h)$. It is specific to a given waveform, detector, event trigger generator and data epoch. All fits resulted in good fit quality, except for the smallest values of h_0, where noise triggers dominate and our detection efficiency is close to 0. Examples are shown in figure 7.

The single-interferometer efficiency curves are then averaged over the antenna pattern and combined to yield the overall efficiency of the burst analysis pipeline to *ad hoc* benchmark bursts. Finally, the efficiency curves are combined with the uninterpreted result and translated into a waveform-dependent bound on the event rate:

$$\mathrm{Rate}_\alpha(h_0) = \frac{N_\alpha}{\varepsilon(h_0) T_{\mathrm{live}}}, \tag{4}$$

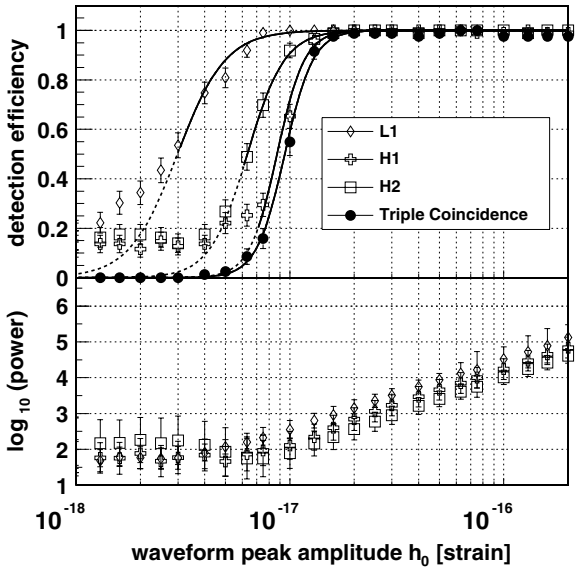

Figure 7. Burst detection efficiency for Gaussian bursts of the form $h(t) = h_0 \exp(-t^2/\tau^2)$ with $\tau = 1$ ms, as a function of peak strain amplitude h_0, for each of the three LIGO detectors and for the triple coincidence, using the TFCLUSTERS event trigger generator. The top graph shows the efficiencies, evaluated through simulations over S1 data. The simulated points are fitted to sigmoid curves in the region where the efficiency is not dominated by random noise triggers. The bottom graph shows the detected power for the simulated bursts. The change of curvature in the power versus strength plots corresponds to the transition between noise and simulated events.

where N_α is the upper limit (with α confidence level) on the number of foreground events, set with the Feldman–Cousins unified approach, and $T_{\rm live}$ is the total livetime.

The result is a set of exclusion plots of *rate versus strength* of gravitational wave bursts originating from fixed strength sources positioned on a fixed sphere centred on the Earth. These plots will be released as soon as the burst analysis is internally reviewed and approved by the entire LIGO Science Collaboration.

A richer interpretation invoking astrophysically motivated signal waveforms as well as source depth and angular distributions are currently under study and will be implemented in future LIGO analysis.

9. Future plans

In the near future we plan to improve our ability to reject random coincidences of events between the interferometers by utilizing the full power of the coincidence analysis. Amplitude information matching and cross-correlation techniques are expected to significantly reduce the number of coincidence events from the analysis pipeline we just described. Within the same time frame, new search algorithms are expected to become available, providing complementary ways of searching for bursts.

The search for transients of gravitational radiation is well under way within the LIGO Science Collaboration. The first science run ('S1') of the LIGO instruments is being followed by an at least ten times more sensitive second science run ('S2') that will last for two months between 14 February and 14 April 2003. Observation during S2 by the LIGO detectors will

be in coincidence with the TAMA detector in Japan, thus providing the opportunity to perform a quadruple coincidence analysis of interferometric data looking for bursts.

Acknowledgment

The LIGO Laboratory gratefully acknowledges the support of the US National Science Foundation under Cooperative Agreement no PHY-0107417.

References

[1] http://www.ligo.caltech.edu
[2] http://www.geo600.uni-hannover.de
[3] http://www.ligo.org
[4] Thorne K S 1995 *Preprint* gr-qc/9506086
[5] http://www.ligo.caltech.edu/∼ajw/bursts/bursts.html
[6] LIGO Scientific Collaboration 2003 *LIGO Internal Document* P030011-00-Z (journal paper in preparation)
[7] http://www.hep.anl.gov/ndk/hypertext/nu_industry.html
[8] Finn L S *et al* 1999 *Phys. Rev.* D **60** 12110
[9] http://www.ldas-cit.ligo.caltech.edu
[10] Daw E *LIGO/LSC Algorithm Library* webpage http://www.lsc-group.phys.uwm.edu/lal
[11] Arnaud N *et al* 1999 *Phys. Rev.* D **59** 82002
[12] Pradier T *et al* 2001 *Phys. Rev.* D **63** 42002
[13] Sylvestre J 2002 *Phys. Rev.* D **66** 102004
[14] Anderson W *et al* 2001 *Phys. Rev.* D **63** 042003
[15] Ito M webpage http://blue.ligo-wa.caltech.edu/gds/dmt/Monitors/glitchMon
[16] Bork R, Shoemaker D, Sigg D and Zweizig J 1999 *Tech. Rep.* LIGO-T990018-A-D
[17] Fritschel P *et al* 2001 *Appl. Opt.* **40** 4988
[18] Feldman G J and Cousins R D 1998 *Phys. Rev.* D **57** 3873