

Using the null-stream of GEO 600 to veto transient events in the detector output

M Hewitson and P Ajith

Max-Planck-Institut für Gravitationsphysik (Albert-Einstein-Institut) und Universität Hannover, Außenstelle Hannover, Callinstr. 38, 30167 Hannover, Germany

E-mail: martin.hewitson@aei.mpg.de

Received 3 August 2005, in final form 29 September 2005

Published 31 October 2005

Online at stacks.iop.org/CQG/22/4903

Abstract

A network of gravitational wave detectors is currently being commissioned around the world. Each of these detectors will search for gravitational waves from various astronomical sources. One of the main searches underway is for un-modelled, transient gravitational wave events. The nature of these signals is such that it will be difficult to distinguish them from bursts of instrumental noise that originate in or around the detector and which then couple to the main detector output. One way to deal with this is to look for events that are coincident in more than one gravitational wave detector. However, with very large event lists (potentially thousands of events per day per detector), the number of events that pass this test due to random chance can still be large. At each detector site, various methods are being developed to veto instrumental bursts from lists of candidate events from that particular detector. This reduces the size of the event lists of each detector, and hopefully the final coincident event list, to a more manageable level. This paper presents one such veto method that can be used to veto certain classes of transient events detected in the output data stream of GEO 600. The method uses events detected in the *null-stream* output of GEO 600 (which contains, in principle, no gravitational wave signal) with a threshold to veto events detected in the main strain output. We show that, for the certain types of signals tested, the method is very robust, delivering high efficiency for a very low false-veto rate. In particular, it is shown that when applied to real detector data, the method is able to strongly veto a certain type of events which appear around 370 Hz in the detector output.

PACS numbers: 95.55.Br, 95.75.Kk, 04.80.Nn, 95.55.Ym

(Some figures in this article are in colour only in the electronic version)

1. Introduction

GEO 600 [1] is an interferometric gravitational wave detector based near Hannover, Germany. The optical layout of GEO 600 is based on a standard Michelson interferometer with the addition of two mirrors: the power-recycling mirror (PR) and the signal-recycling mirror (SR). This so-called *dual-recycled* optical scheme is used to enhance the sensitivity of the detector (see [2] and [3] for details).

The use of signal-recycling and a heterodyne readout scheme means that gravitational wave signals can only be extracted optimally from the detector output at one frequency by selecting an appropriate demodulation phase of the output signal. This means that, for any other frequency, the gravitational wave information will be spread between two orthogonal demodulation quadratures of the output.

In GEO 600, both output quadratures are calibrated to strain since, in general, the noise components of each output quadrature are different. This calibration is done in the time domain by injecting calibration lines into the differential length-control actuators of the Michelson interferometer. These calibration lines produce, at least at those frequencies, known strain signals. From the measurement of these calibration lines in the detector output, the response function of the detector is determined periodically on-line. The inverse of this response function is used to compute time-domain filters which are used to filter the detector outputs, $P(t)$ and $Q(t)$, to produce two estimates of the detected strain, $h_P(t)$ and $h_Q(t)$. Details of the calibration method used at GEO 600 are given in [4, 5].

If we consider that both $h_P(t)$ and $h_Q(t)$ contain the same gravitational wave signal, $h(t)$, together with different noise components, $N_P(t)$ and $N_Q(t)$, such that

$$h_P(t) = h(t) + N_P(t), \quad (1)$$

$$h_Q(t) = h(t) + N_Q(t), \quad (2)$$

then we can recover the best estimate of the underlying signal, $h(t)$, using a maximum likelihood method to combine $h_P(t)$ and $h_Q(t)$. Details of this combining method are given in [6].

We can also form the *null-stream* from $h_P(t)$ and $h_Q(t)$ by

$$h_{\text{null}}(t) = h_P(t) - h_Q(t), \quad (3)$$

which will yield a signal that contains no gravitational wave signal to a level consistent with the relative calibration accuracy of $h_P(t)$ and $h_Q(t)$.

Figure 1 shows snapshot amplitude spectral density estimates of both $h(t)$ and $h_{\text{null}}(t)$ from the same stretch of data; the injected calibration lines are highlighted. We can see that the calibration lines (and hence any gravitational wave signal), appear significantly suppressed in the null-stream. We can also see that since the amplitude of the calibration lines is not zero in the null-stream, there is a relative error in the calibration of the two output quadratures of the order of 5% over most of the frequency band.

We can search for signals in this null-stream, as we do in the $h(t)$ signal. If a coincident signal is found in both the null-stream and the $h(t)$ stream, with a relative amplitude that is *inconsistent* with it being a gravitational wave signal, then we can veto it. This veto method is studied in the following sections using software injections into real data from GEO 600. A false-veto rate and efficiency of the veto are developed as a function of an amplitude consistency threshold. This method is then applied with a particular threshold to a long stretch of GEO 600 data.

Several veto methods have been proposed and implemented in the search for gravitational waves from different kinds of sources. One popular method is to look for transient events in

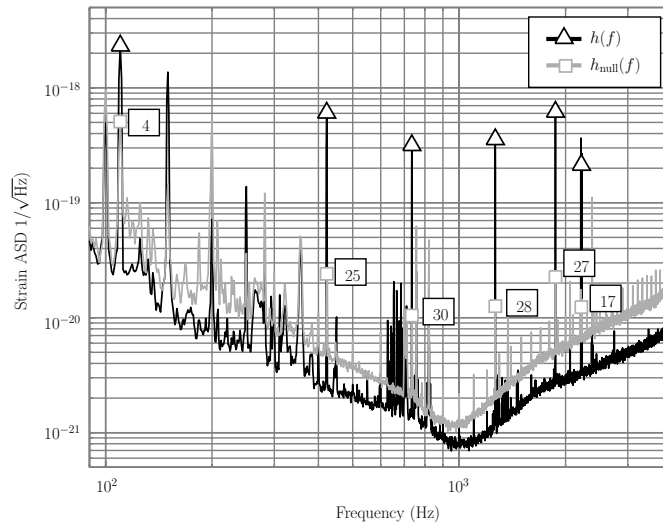


Figure 1. Amplitude spectral densities of the strain and null-stream outputs of GEO 600. Calibration lines are marked for the two spectra, and the ratio of the calibration line amplitudes in the strain signal and the null-stream signal is indicated in the boxes.

the auxiliary and environmental channels of the detector and veto those events which couple with the main detector output. A summary of the recent work in this area can be found in [7–11]. Another class of statistical vetoes have been developed which rely on specific models of waveforms; see [12–16] for details. These statistical vetoes are very useful in the search for gravitational waves from pulsars and inspiralling compact binaries, where accurate models of the expected waveforms are known. While the veto strategy proposed in this paper relies on the physical properties of the gravitational waves, it does not make any assumptions about the model of the waveforms and is, in this sense, very robust. Another null-stream method for vetoing burst events that are coincident in multiple detectors (which is based on the formulation laid out in [17]), is also under development.

2. Signal injection and detection

In order to study the false-veto rate and efficiency of such a veto method, we used software injections into the two detector outputs, $P(t)$ and $Q(t)$, prior to the calibration routine being applied. The signals injected were sine-Gaussians of the form

$$s(t) = h_{\text{rss}} \left(\frac{2f_0^2}{\pi} \right)^{1/4} \sin[2\pi f_0(t - t_0)] \exp \left[-\frac{(t - t_0)^2}{2\tau^2} \right], \quad (4)$$

where h_{rss} is a scaling amplitude such that the total integrated power in the waveform is equal to h_{rss}^2 , f_0 is the central frequency of the waveform, τ sets the duration of the signal, and an offset of $t_0 = 0.5$ s is included just to make the waveform peak in the middle of a 1 s data segment. An additional random offset from the second boundary, T_{off} , is also included at the point of injection.

In order to detect these signals in both the strain signal and the null-stream, we used the HACR algorithm (hierarchical algorithm for clusters and ridges) to generate lists of transient events. The details of this algorithm can be found in [18] and [19]. The algorithm essentially looks for short duration excesses of power in each frequency bin of a time–frequency map

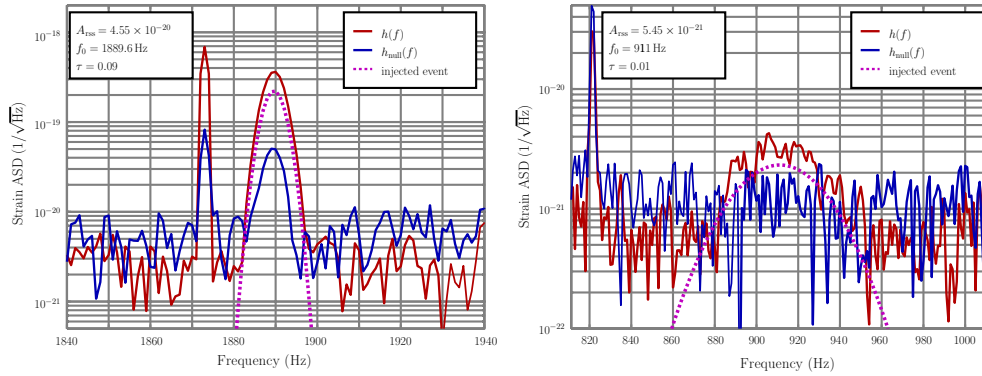


Figure 2. Two example injected events into real data. The amplitude spectral density estimates are computed from 1 s of data with no averaging and using a Hanning window. The parameters of the injected waveform are shown in the legend. The left plot shows an event with a high SNR; the right plot, one with a low SNR. Each plot contains a large line: in the left plot, it is an injected calibration line; in the right plot, it is a seismically induced line from a turbo pump that is part of the vacuum system.

Table 1. The ranges of the waveform parameters used in the series of software injection experiments.

Parameter	Min	Max
h_{rss}	5×10^{-21}	5×10^{-20}
f_0 (Hz)	900	2000
τ (s)	0.01	0.1
T_{off} (samples)	0	f_s

of the data. Neighbouring pixels that show excess power are clustered together into events. Some characteristics of each event (time, central frequency, total power, etc) are estimated and logged. This algorithm uses an adaptive threshold for trigger generation (set based on the particular data segment being analysed).

In all of the experiments detailed below (unless stated otherwise), the parameters of the sine-Gaussian injections were allowed to vary in the ranges shown in table 1. These parameter ranges lead to injected signals that span a sensible range of signal-to-noise ratio (SNR) in the detector strain signal: not too small as to be undetectable by HACR, and not too big as to be unrealistic. Figure 2 shows two examples of injected events; one showing a typically small SNR injection, and one showing a typically large SNR injection.

We only apply the veto to events which are coincident events in $h(t)$ and $h_{\text{null}}(t)$. These coincident events are selected from the two event lists by applying time- and frequency-consistency windows, Δt and Δf respectively. The size of each of these windows was selected using software injections into the strain signal. Figure 3 shows efficiency curves for these two tunable parameters, Δt and Δf , generated by comparing the detected events to the injected events from a particular set of 200 random injections using the parameter ranges indicated in table 1. These events were injected into the two uncalibrated detector output streams with relative amplitudes such that the amplitudes in the two calibrated output streams, $h_P(t)$ and $h_Q(t)$, are equal (yielding only one detected event per injection). (This can also be done by injecting into only one of the uncalibrated output data streams.) From these injections, 193 were detected by HACR in the $h(t)$ stream.

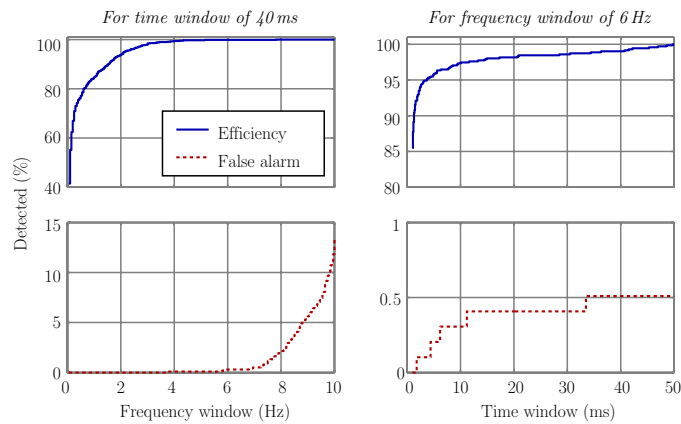


Figure 3. False alarm and efficiency plots for the time and frequency cuts (Δt and Δf) applied to events detected by HACR. From these plots, a frequency window of 6 Hz and a time window of 40 ms were chosen for the null-stream experiments. Here ‘efficiency’ refers to the percentage of coincident events that are expected to pass the consistency test; ‘false alarm’ refers to the percentage of non-coincident events which are falsely taken to be coincident when applying the consistency test.

The curves shown in figure 3 give the false-alarm rate and efficiency of the time and frequency selection of these 193 detected events. The false-alarm rate curve for each parameter is generated by time or frequency shifting the detected events by 1.1 s and 10.2 Hz respectively prior to making the time and frequency cut; those that pass the consistency test are considered a false alarm. From this experiment, we chose a time-consistency window of 40 ms and a frequency consistency of 6 Hz to apply to the $h(t)$ and $h_{\text{null}}(t)$ event lists when using the HACR algorithm in the experiments detailed in the rest of this paper.

3. The amplitude consistency test

The amplitude consistency test that is applied to the event lists is a simple threshold test. We form the ratio of the amplitude of events that are time and frequency consistent in the two event lists (from $h(t)$ and $h_{\text{null}}(t)$). If the ratio, h/h_{null} , of a particular event is below a certain threshold, A_{thresh} , then we can veto the event.

The value of the threshold, A_{thresh} , can be determined in two ways: a loose condition can be set by observing the amplitude of the calibration lines in both $h(f)$ and $h_{\text{null}}(f)$. Figure 1 shows the ratio of the calibration line amplitudes in $h(f)$ and $h_{\text{null}}(f)$ as a function of frequency. From this we could set a threshold of around 25—the mean of the ratios. This is only one way in which we can set the threshold. We can determine a much more stringent test by performing software injections and looking at the efficiency of the veto and the false-veto rate as a function of the threshold we apply.

3.1. False-veto rate and efficiency studies

Four different sets of 200 sine-Gaussian waveforms were injected into 2000 s of the two uncalibrated detector output data streams of GEO 600 in order to perform the four experiments detailed in the following subsections. The time interval between injections was allowed to vary (randomly) between 0 and 10 s. The waveform parameters were randomly chosen from

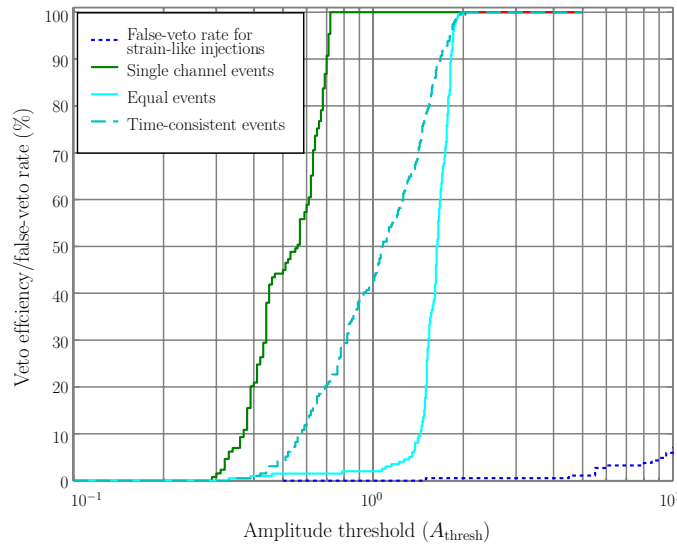


Figure 4. The false veto and efficiency curves for the null-stream veto. The efficiency curves are computed for different cases of signal injection—see text for details.

the ranges given in table 1. The HACR algorithm was then used to detect events in both $h(t)$ and $h_{\text{null}}(t)$.

In all four experiments, a time and frequency consistency cut (using the windows discussed above) was applied to these two event lists to preferentially select the injected events. These two shorter lists were then tested for amplitude consistency for various values of A_{thresh} .

Strain-like injections. For this experiment, the waveforms were injected with relative amplitudes consistent with the signals from a gravitational wave source. In other words, the injections into the two *uncalibrated* data streams had relative amplitudes such that they will have equal amplitude in the *calibrated* $h_P(t)$ and $h_Q(t)$ strain outputs. Typical detector response functions for P and Q were used to generate the signals to add to the two uncalibrated data streams. These detector response functions are not necessarily those used to calibrate the $P(t)$ and $Q(t)$ time series, thus mimicking the possible effects of calibration errors. In order to detect the injections in the null-stream, the range of amplitudes of the injections was increased by a factor 5. From the 200 injections, 197 were detected in the $h(t)$ stream, and 99 were detected in the $h_{\text{null}}(t)$ stream; from these two lists, 99 events were found to be time and frequency consistent.

Figure 4 shows the result of applying the veto to the two time and frequency consistent events lists as a function of threshold. Since we expect the vast majority of the events in the time–frequency consistent lists to be GW-like events, any events that are vetoed are attributed to the false-veto rate of the method for that particular threshold setting.

Single output-channel events. One test of the efficiency of the veto method can be done by considering injections into the two uncalibrated detector outputs that are discrete in time, i.e., injected into either detector output, but not both. This simulates those real-life events that couple into the two detector outputs after the demodulation process where the two signal paths are separated. Of the 200 injected events, 169 were detected in the $h(t)$ stream, and 178 were detected in the null-stream. Of the events of these two lists, 165 were found to be time and frequency consistent.

The efficiency curve for this class of events, generated from these two event lists, is labelled ‘Single channel events’ in figure 4. Such events always appear in the null-stream with a higher amplitude than in the $h(t)$ stream since the null-stream contains all the signal whereas the $h(t)$ stream gets a share of the power from both calibrated strain outputs (see [6] for details). Because of this, we expect these signals to be vetoed with a very high efficiency.

Equal amplitude events. This class of injections are simultaneous and have equal amplitude in both of the *uncalibrated* detector output signals. Such events could arise due to a coupling in the data acquisition system where a third channel couples transient signals into both detector output signals prior to digitization. Of the 200 injected events, 189 were detected in the $h(t)$ stream, and 182 were detected in the null-stream. Of the events of these two lists, 172 were found to be time and frequency consistent.

The efficiency curve for this type of events is labelled ‘Equal events’ in figure 4. Events such as these should appear slightly stronger in $h(t)$ stream than in the null-stream since the signal in the null-stream is reduced by the ratio of the two optical transfer functions at the frequency of the event. Since this ratio is, for most frequencies, of the order 2 or less (see [5] for typical optical transfer functions of GEO 600), we would expect to efficiently veto them.

Random amplitude, simultaneous events. The final class of signals studied are simultaneous in both uncalibrated detector outputs, but have a random relative amplitude in each. For these injections, a single waveform was selected at random using the parameter ranges given in table 1. Two copies of this waveform were then injected in the two uncalibrated detector output signals, with each being multiplied by a different random scaling factor selected from the range 0.1 to 10 prior to injection. This type of signal would be symptomatic of a noise source that couples at the detector output, either before or after the demodulation process, but in a way that means the ratio of the two coupling paths is arbitrary and may depend on other external influences. One example of this may be pickup that depends on cable positions or signal sizes. Of the 200 injected events, 175 were detected in the $h(t)$ stream, and 178 were detected in the null-stream. Of the events of these two lists, 160 were found to be time and frequency consistent.

The efficiency curve for this type of event is labelled ‘Time-consistent events’ in figure 4.

3.2. Setting the consistency threshold

From figure 4, we can safely set a threshold for this veto method of $A_{\text{thresh}} = 5$. The threshold was chosen to be as high as possible (to maximize the efficiency) for a false-alarm rate of less than 1%. From this figure, we can see that we expect to get a false-veto rate of around 1% and an efficiency (at least for the type of signals explored so far) of almost 100%.

4. Application to GEO data

4.1. Extended section of science data

The veto method was applied to a long stretch (22 h) of calibrated data from GEO 600. For the entire duration, the detector was locked and operating normally.

The HACR algorithm was run on the $h(t)$ and $h_{\text{null}}(t)$ data streams for this time. A total of 141 766 events were found in $h(t)$, 42 380 in $h_{\text{null}}(t)$. Of these, 1086 events were found to be time and frequency consistent using the consistency windows described above.

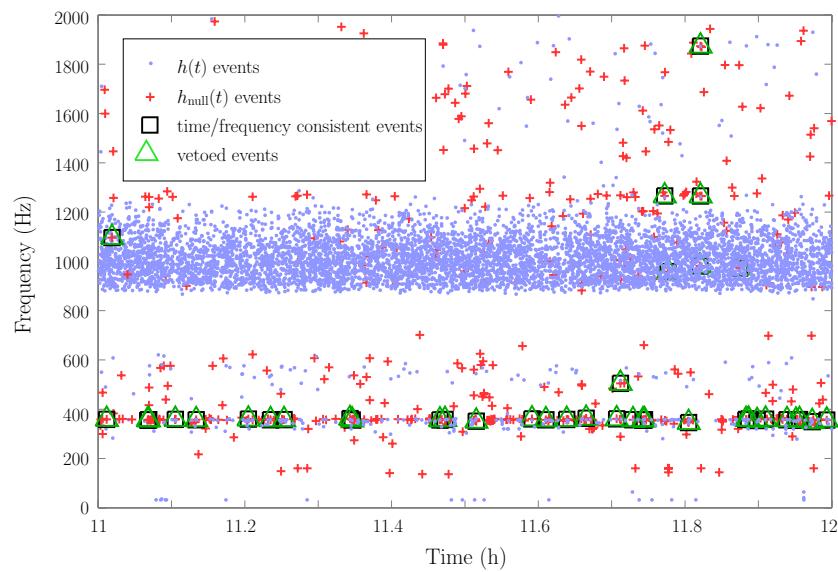


Figure 5. A time–frequency map of the events detected by HACR in $h(t)$ and $h_{\text{null}}(t)$ for a 1 h data stretch of GEO 600. Those events that were found to be time and frequency consistent between the two lists are marked with boxes; those that were subsequently vetoed are marked with triangles.

The amplitude consistency veto was applied to these two reduced event lists with a threshold of 5. From this we would expect to falsely veto around 10 events. All 1086 of the time and frequency consistent events were found to be inconsistent in amplitude for being a gravitational wave, and were hence vetoed. Figure 5 shows a short section of the events from the 22 h of data.

We can see immediately that the majority of vetoed events have a central frequency around 370 Hz; this is true for the entire 22 h data stretch. At this frequency, the output noise of the detector is dominated by features which are believed to be acoustically or seismically driven resonances of some of the optical mounts that steer the output beam to the main output photodetector. For such events, the veto method seems to be very efficient, as the software injections would suggest.

The other (few) events that are vetoed could be false vetoes, although the number of vetoes outwith the 370 Hz band suggests that at least some of them are not. In addition, the fact that a large concentration of the events around 1 kHz are detected in $h(t)$ but not in $h_{\text{null}}(t)$ suggests that these instrumental bursts pass through the detector in a similar way to gravitational waves and hence appear with similar amplitudes in the two calibrated output streams (most probably frequency noise originating in the laser subsystem).

5. Summary

The null-stream output of GEO 600 has been used to generate a list of transient events that can be used to veto those events detected in the main strain output of GEO 600. By using software injections, an amplitude-ratio threshold has been determined that can be used to compare the amplitude of events in the $h(t)$ data stream to those events detected in the null-stream, $h_{\text{null}}(t)$. Events that were found to be time and frequency consistent in the two event lists could be vetoed with 100% efficiency for a false-veto rate of 1%. In addition, this method was applied

to a short section of real detector data and it was possible to veto a significant fraction of events that appeared clustered around 370 Hz in the detector output.

One possible way to improve the performance of this method would be to have a threshold that is a function of frequency. Currently, the relative calibration accuracy of the two detector outputs is different for different frequencies, varying from about 20% at low frequencies, to around 5% at high frequencies. This means that the suppression of gravitational wave events in the null-stream, and hence the threshold needed to robustly test them, is also a function of frequency.

While the results presented in the paper are based on the trigger events produced by the HACR algorithm, the principles should apply for any algorithm used to generate the event lists. However, it is clear that the chosen amplitude consistency threshold may be different when different trigger generation algorithms are used. This remains to be tested. In addition, the results presented are based only on software injections. It remains to do some tests using hardware injections (bursts injected at various physical points in the detector subsystems) to provide confidence in the method.

Acknowledgments

The authors would like to thank the entire GEO 600 team for their work in producing the GEO 600 detector which makes the study presented in this paper both relevant and possible. We are also grateful for support from PPARC and the University of Glasgow in the UK, and the BMBF and the state of Lower Saxony in Germany.

References

- [1] Smith J R *et al* 2004 Commissioning, characterization, and operation of the dual-recycled GEO 600 *Class. Quantum Grav.* **21** S1737–45
- [2] Heinzel G *et al* 2002 Dual recycling for GEO 600 *Class. Quantum Grav.* **19** 1547
- [3] Grote H, Freise A, Malec M, Heinzel G, Willke B, Lück H, Strain K A, Hough J and Danzmann K 2004 Dual recycling for GEO 600 *Class. Quantum Grav.* **21** S473–80
- [4] Hewitson M, Grote H, Heinzel G, Strain K A, Ward H and Weiland U 2003 Calibration of the power-recycled gravitational wave detector, GEO600 *Rev. Sci. Instrum.* **74** 4184
- [5] Hewitson M, Heinzel G, Smith J R, Strain K A and Ward H 2004 Principles of calibrating the dual-recycled GEO 600 *Rev. Sci. Instrum.* **75** 4702
- [6] Hewitson M, Grote H, Hild S, Lück H, Ajith P, Smith J R, Strain K A, Willke B and Woan G Optimal time-domain combination of the two calibrated output quadratures of GEO 600 *Class. Quantum Grav.* **22** 4253
- [7] Kotter K, Heng I S, Hewitson M, Strain K A, Woan G and Ward H 2003 PQMon: a powerful veto for burst events *Class. Quantum Grav.* **20** S895–902
- [8] Christensen N, Shawhan P and González G (for the LIGO Scientific Collaboration) 2004 Vetoes for inspiral triggers in LIGO data *Class. Quantum Grav.* **21** S1747–55
- [9] Christensen N (for the LIGO Scientific Collaboration) 2005 Veto studies for LIGO inspiral triggers *Class. Quantum Grav.* **22** S1059–68
- [10] Di Credico A (for the LIGO Scientific Collaboration) 2005 Gravitational wave burst vetoes in the LIGO S2 and S3 data analyses *Class. Quantum Grav.* **22** S1051–8
- [11] Ajith P, Hewitson M and Smith J R *et al* 2005 Instrumental vetoes for gravitational wave burst triggers using noise projections, in preparation
- [12] Baggio L *et al* 2000 χ^2 testing of optimal filters for gravitational wave signals: an experimental implementation *Phys. Rev. D* **61** 102001
- [13] Itoh Y, Papa M A, Krishnan B and Siemens X 2004 Chi-square test on candidate events from CW signal coherent searches *Class. Quantum Grav.* **21** S1667–77
- [14] Shawhan P and Ochsner E 2004 A new waveform consistency test for gravitational wave inspiral searches *Class. Quantum Grav.* **21** S1757–65

-
- [15] Allen B 2005 χ^2 time-frequency discriminator for gravitational wave detection *Phys. Rev. D* **71** 062001
 - [16] Babak S, Grote H, Hewitson M, Lück H and Strain K A 2005 Signal based vetoes for the detection of gravitational waves from inspiralling compact binaries *Phys. Rev. D* **72** 022002
 - [17] Gürsel Y and Tinto M 1989 Near optimal solution to the inverse problem for gravitational-wave bursts *Phys. Rev. D* **40** 3884–938
 - [18] Heng I S, Balasubramanian R, Sathyaprakash B S and Schutz B F 2004 First steps towards characterizing the hierarchical algorithm for curves and ridges pipeline *Class. Quantum Grav.* **21** S821
 - [19] Balasubramanian R, Grote H, Heng I S, Hewitson M, Lück H, Smith J R, Strain K A, Ward H and Willke B 2005 Results from the first burst hardware injections performed on GEO 600 *Class. Quantum Grav.* **22** 3015–28