A first comparison of search methods for gravitational wave bursts using LIGO and Virgo simulated data

F Beauville[1], M-A Bizouard[2], L Blackburn[3], L Bosi[4], P Brady[5], L Brocco[6], D Brown[5,7], D Buskulic[1], S Chatterji[7], N Christensen[8], A-C Clapson[2], S Fairhurst[5], D Grosjean[1], G Guidi[9,10], P Hello[2], E Katsavounidis[3], M Knight[8], A Lazzarini[7], F Marion[1], B Mours[1], F Ricci[6], A Viceré[9,10] and M Zanolin[3] (The Joint LIGO/Virgo Working Group)

[1] Laboratoire d'Annecy-le-Vieux de physique des particules, Chemin de Bellevue, BP 110, 74941 Annecy-le-Vieux Cedex, France
[2] Laboratoire de l'Accélérateur Linéaire (LAL), IN2P3/CNRS-Université de Paris-Sud, BP 34, 91898 Orsay Cedex, France
[3] LIGO-Massachusetts Institute of Technology, Cambridge, MA 02139, USA
[4] INFN Sezione di Perugia and/or Università di Perugia, Via A Pascoli, I-06123 Perugia, Italy
[5] University of Wisconsin–Milwaukee, Milwaukee, WI 53201, USA
[6] INFN, Sezione di Roma and/or Università 'La Sapienza', P.le A. Moro 2, I-00185 Roma, Italy
[7] LIGO-California Institute of Technology, Pasadena, CA 91125, USA
[8] Carleton College, Northfield, MN 55057, USA
[9] INFN-Sezione Firenze/Urbino Via G Sansone 1, I-50019 Sesto Fiorentino, Italy
[10] Università di Urbino, Via S Chiara 27, I-61029 Urbino, Italy

E-mail: zanolin@mit.edu

Abstract

We present a comparative study of six search methods for gravitational wave bursts using simulated LIGO and Virgo noise data. The simulated data were generated according to the design sensitivity of the two 4 km LIGO interferometers and the 3 km Virgo interferometer. The searches were applied on replicas of the data sets to which eight different signals were injected. Three figures of merit were employed in this analysis: (a) receiver operator characteristic curves, (b) necessary signal-to-noise ratios for the searches to achieve 50% and 90% efficiencies and (c) variance and bias for the estimation of the arrival time of a gravitational wave burst.

PACS numbers: 04.80.Nn, 07.05.Kf, 02.70.Hm, 95.55.Ym

1. Introduction

Progress in the commissioning of the LIGO and Virgo detectors makes possible in the near future the opportunity of a joint network analysis. To prepare the ground, gain a better

understanding of each other's analyses and develop common procedures, the LIGO Scientific Collaboration and the Virgo experiment have agreed to pursue a joint search for burst and binary inspiral signals on simulated data [1]. This paper describes the application of burst search algorithms on simulated data to which several classes of candidate signals waveforms were injected. This framework provides the opportunity to (a) compare the performance of time domain and frequency domain methods, (b) compare their performance over different detectors, (c) investigate the limits in detection at design sensitivity and (d) compare algorithms on the generation of triggers for single interferometers before temporal coincidences are applied.

Three hours of simulated data have been generated with a spectrum following the target design sensitivity of both the 4 km LIGO interferometers, with sampling frequency $f_s = 16\,384$ Hz, and the 3 km Virgo interferometer, with $f_s = 20\,000$ Hz (see the inspiral proceedings in this issue for more details [2]). We only recall here that the simulated noise is a correct representation of the design noise for frequencies above 50 Hz. Below that frequency both the LIGO and Virgo noise depart from a realistic model. Moreover, they depart in a different way because the Virgo data contains a large line around 5 Hz that requires for some methods a preprocessing involving high pass and line removal filters.

Three families of signals have been generated and injected onto the simulated noise with a mean rate of one every 60 s. Sine Gaussian (SG) and Gaussian (GA) injections (see, for example, [3]) were chosen to represent the two general classes of short-lived gravitational wave bursts of narrow-band and broad-band characters, respectively. Sine Gaussian injections were chosen with central frequencies $f = 235$ Hz, to probe the best sensitivity region of the spectrum, and 820 Hz, to probe the higher frequency regime. Both frequencies were tested with quality factors $Q = 5$ and 15. GA signals were chosen with duration equal to 1 and 4 ms. Dimmelmeier–Font–Mueller (DFM) supernovae core collapse signals [4] with parameters $a = 1$, $b = 2$, $g = 1$ and $a = 2$, $b = 4$, $g = 1$ were adopted as waveforms with more complicated time–frequency structure and non-minimal time–frequency area. The specific choice of DFM waveforms was motivated by their different central frequency, and that the type I (a1b2g1) signals typically present a first peak followed by a ringdown, while type II signals (a2b4g1) show a few decreasing peaks.

The strength of a signal, similarly to [3], was quantified by the signal-to-noise ratio (ρ) used to characterize the performance of an optimal filter

$$\rho = \sqrt{4 \int_0^\infty \frac{|h(f)|^2}{\sigma(f)}\,\mathrm{d}f} \tag{1}$$

where $h(f)$ is the Fourier transform of the injected waveform and $\sigma(f)$ is the one-sided noise power spectral density. Injections were performed with ρ from 2 to 10, which generally corresponded to 0–100% detection efficiencies.

The various search methods can be divided into two categories. The first category searches for excess power in the time–frequency domain:

- *Q-transform (QT)* [5]. A multiresolution time–frequency search for excess power applied on data that are first whitened using a zero-phase linear prediction. Equivalent to optimal matched filter for minimum uncertainty waveforms of unknown phase in the whitened data.
- *S-transform (ST)* [6]. A search for statistically significant clusters of power in the time–frequency map generated using a kernel composed of complex exponentials shaped by Gaussian profiles with width inversely proportional to frequency. The ST applies a line removal on Virgo data and a high pass filter on LIGO data.

- *Power filter (PF)* [7]. A search on whitened data for excess power over different time intervals and set of frequencies.

And the second searches for events in the time domain only:

- *Peak correlator (PC)* [8]. A search for peaks of Wiener filtered data with Gaussian templates. PC applies on Virgo data a high pass filter and a line removal filter.
- *Mean filter (MF)* [9]. A search for excess power in moving averages of whitened data over intervals containing from 10 to 200 samples. The MF whitens the data and applies to Virgo data a high pass filter and a line removal.
- *Adaptive linear filter (ALF)* [10]. A search for change in slope over moving windows of data over intervals containing from 10 to 300 samples. ALF applies the same whitening high pass filtering and line removal as MF.

The performance of the searches has been investigated by computing (a) the receiver operator characteristic curves (ROCs) for each method, injected waveform, interferometer and ρ, (b) computing the necessary ρ for a search to reach a detection efficiency equal to 50% and 90% at a particular false alarm rate (FAR) and (c) assessing the accuracy in the estimate of the arrival peak time of a signal.

In section 2, we describe the analysis pipeline. In section 3, we discuss a selection of the figures of merit generated to compare the search methods. In section 4, we summarize the findings of this project and address future directions of the investigation.

2. Analysis pipeline

Simulated strain data corresponding to noise and signal from both instruments were made available in the frame format to both the collaborations. Each method was then applied to noise plus signals for FARs between 10^{-4} Hz, which corresponds to roughly one false trigger in 3 h, and 0.1 Hz, which allowed us to study the statistical distribution of durations and temporal separation of the triggers. The performance of the searches at higher single-interferometer FARs is also relevant for network studies where consistency criteria can be applied on the triggers giving a lower FAR. For example, requiring arrival time coincidence alone dramatically lowers the FAR in a multi-IFO analysis, and other tests such as frequency and amplitude consistency can also be performed.

Each method computed a start time, a peak time and an end time for each of the events, as well as the efficiencies in the chosen range of FARs. The searches adopted different procedures to associate triggers with injections and therefore compute the efficiencies. However, methods that use matching windows to identify detected injections with triggers can freely choose the size of the matching window provided that it is (I) larger than the search time resolution and (II) the products of the matching window for the false alarm rate and the injection rate are both much smaller than 1. In this analysis: (a) the PF associated with each injected waveform a time interval starting at the first non-zero sample of the simulated waveform that is added on the data and finishing at the last non-zero sample. A detection was claimed when the trigger interval, as computed by the PF, overlapped with the event interval. (b) For the QT, the interval internally associated with a trigger needed to overlap with a 0.2 s interval centred at the injection peak time. (c) For MF, ALF, PC and ST the time associated with a trigger needed to be closer to the peak time of the injection than 20 ms for ALF, MF and ST, and 50 ms for PC. We will show in section 3 that constraint (I) was considered in the analysis. Constraint (II) was also considered. In fact, the largest FAR that we study is 0.1 Hz and the largest coincidence window is 0.05 s while the injection rate is 1/60 Hz. This makes the two products of point (II) both smaller than at least 0.01.

Table 1. Necessary ρ for 0.5 and 0.9 efficiencies at FAR = 0.01 Hz for LIGO data.

LIGO efficiency	MF 0.5	ALF	PC	PF	ST	QT	MF 0.9	ALF	PC	PF	ST	QT
SG235q5	7.5	6.4	6.1	6.8	5.2	5.1	10.2	8.2	8.4	8.8	6.3	6.4
SG235q15	10.9	9.2	NA	6.3	5.4	5.1	15.5	12.6	NA	8.5	6.7	6.5
SG820q5	8.7	7.4	NA	7.6	NA	5.2	11.7	9.5	NA	9.9	NA	6.6
SG820q15	15.4	11.1	NA	6.8	NA	5.1	24.4	15.4	NA	8.5	NA	6.3
GA1d0	6.6	6.3	4.9	7.1	5.5	5.5	8.6	7.9	6.2	9.1	6.8	6.7
GA4d0	7.5	6.4	4.9	6.8	5.9	5.6	9.6	8.1	6.0	8.6	7.4	7.0
DFMa1b2g1	7.3	6.7	5.1	7.8	6.9	6.5	9.3	8.6	6.4	10.1	8.4	8.2
DFMa2b4g1	6.9	6.6	5.8	8.0	6.6	6.4	9.2	8.6	7.5	10.3	8.2	8.0

Table 2. Necessary ρ for 0.5 and 0.9 efficiencies at FAR = 0.01 Hz for Virgo data.

Virgo efficiency	MF 0.5	ALF	PC	PF	ST	QT	MF 0.9	ALF	PC	PF	ST	QT
SG235q5	7.5	6.9	6.6	6.5	5.7	5.2	9.9	8.7	8.7	8.4	7.2	6.4
SG235q15	11.1	9.6	NA	6.4	5.6	5.1	16.5	12.6	NA	8.4	7.0	6.6
SG820q5	8.1	6.8	NA	7.6	NA	5.1	10.6	8.4	NA	9.6	NA	6.5
SG820q15	19.1	9.9	NA	7.3	NA	5.3	23.8	13.6	NA	9.4	NA	6.8
GA1d0	5.8	5.6	4.9	7.2	5.6	5.9	7.6	7.1	6.0	9.1	7.0	7.5
GA4d0	6.2	5.3	5.2	8.7	5.4	7.0	7.8	6.7	6.1	11.7	6.8	9.1
DFMa1b2g1	6.1	5.9	5.6	7.7	5.7	6.6	7.7	7.5	6.9	9.4	7.2	8.2
DFMa2b4g1	6.3	6.0	5.8	8.3	5.5	6.8	8.1	7.6	7.5	10.9	6.9	8.8

3. Data analysis

A key element in comparing search methods are ROCs. Each point of a ROC curve represents the efficiency in detecting a certain waveform for a given false alarm rate. Different points are obtained for different values of internal parameters of the different search methods. For a given waveform, ROCs allow us to compare different methods. Ideally, one would like to generate ROCs curves for a family of signals that are representative of all the possible observable signals and test the robustness of the prediction versus small variation of the waveform properties. In this preliminary study, we simply present the ROCs for a choice of signals with different duration, bandwidth and profiles.

In a single-interferometer experiment the FAR could be chosen by considering much less than one noise event over the whole acquisition time in order to achieve confidence that an observed event is a true gravitational wave event. More generally, the largest tolerable single-interferometer FAR will depend on the observation time, number of interferometers and consistency requirements between triggers generated from the data belonging to different interferometers. In this perspective, as informative figures of merit, we first present examples of ROC curves and the necessary ρ to reach 50% and 90% detection efficiencies for FAR = 0.01 Hz. Of all the possible parameters on which consistency constraints can be applied, in this study, we focused on the trigger peak time, which should be compatible with the travel time between interferometers and the method errors in the peak time estimation. In tables 1–4, some results corresponding to the PC and ST running on SG are not available. They reflect waveforms where the detection efficiency of the PC and ST was not significant.

We present representative ROC curves for SG235Q5 and DFMa1b2g1, two values of ρ (10 and 5), on Virgo data in figures 1 and 2, and the ROC curves for SG820Q15 and GA1d0, two values of ρ (10 and 5), on LIGO data in figures 3 and 4. ROCs for the same waveform and different interferometers are not shown since the performance of the methods

Table 3. Standard deviation in milliseconds for peak time estimation $\rho = 10$ and FAR $= 0.1$ Hz.

Standard deviation	MF	ALF	PC	PF	ST	QT
SG235q5 LIGO	1.4	1.2	1.2	21.7	0.7	0.7
SG235q5 Virgo	1.5	1.3	1.3	13.8	0.8	0.9
SG235q15 LIGO	4.1	4.2	NA	15.2	1.7	1.9
SG235q15 Virgo	4.9	5.2	NA	11.7	2.0	2.5
SG820q5 LIGO	0.4	0.4	NA	15.9	NA	0.2
SG820q5 Virgo	0.4	0.3	NA	16.9	NA	0.2
SG820q15 LIGO	1.3	1.3	NA	19.1	NA	0.6
SG820q15 Virgo	8.4	3.3	NA	14.5	NA	0.7
GA1d0 LIGO	1.2	1.6	0.1	15.5	0.7	0.7
GA1d0 Virgo	1.2	0.6	0.1	16.4	0.8	0.8
GA4d0 LIGO	1.6	3.5	0.3	15.5	5.8	1.5
GA4d0 Virgo	2.5	1.6	0.3	12.3	1.4	1.9
DFMa1b2g1 LIGO	0.3	1.0	0	23.0	1.4	0.7
DFMa1b2g1 Virgo	0.3	0	0	15.2	0.4	0.2
DFMa2b4g1 LIGO	1.7	2.8	1.0	18.1	2.6	2.2
DFMa2b4g1 Virgo	2.5	0.3	0.2	18.1	2.5	1.4

Table 4. Bias in milliseconds for peak time estimation $\rho = 10$ and FAR $= 0.1$ Hz.

Bias	MF	ALF	PC	PF	ST	QT
SG235q5 LIGO	1.8	0.2	0.2	5.0	1.7	−0.1
SG235q5 Virgo	0.7	−0.5	−0.5	3.9	0.4	−0.1
SG235q15 LIGO	1.9	0.3	NA	5.8	1.5	−0.2
SG235q15 Virgo	0.6	−0.1	NA	9.5	0.5	0
SG820q5 LIGO	0.1	−0.4	NA	7.2	NA	0
SG820q5 Virgo	0.1	−0.4	NA	5.1	NA	0
SG820q15 LIGO	0.3	−0.4	NA	4.8	NA	−0.1
SG820q15 Virgo	1.1	0	NA	5.1	NA	0
GA1d0 LIGO	1.2	3.3	0	0.4	3.7	0
GA1d0 Virgo	−0.7	2.3	0	−0.7	1.6	−0.1
GA4d0 LIGO	3.3	8.0	0	−4.4	2.9	0
GA4d0 Virgo	−0.2	3.3	0	3.6	0	−0.2
DFMa1b2g1 LIGO	−0.2	0.6	−0.1	8.6	1.5	0.2
DFMa1b2g1 Virgo	−0.4	0	−0.1	6.1	0	−0.1
DFMa2b4g1 LIGO	2.4	4.4	−1.0	−4.6	4.6	−0.9
DFMa2b4g1 Virgo	−2.8	2.0	−1.3	1.0	2.6	−1.6

gave fairly similar results over the two noise spectra. Time domain methods were observed to perform better on GA pulses and DFM supernovae core collapse while time–frequency domain methods had a better performance on the SG signals. The two families of curves corresponding to the two values of ρ of the injected signals also illustrate how the algorithm performance degrades with decreasing ρ. SG results follow intuition since the time–frequency tailoring of the QT and ST is close to a matched filter for minimum uncertainty waveforms of unknown phase in the whitened data. A similar line of thought applies to PC which is an optimal filter for GA pulses. The performance of a search is often quantified through the conditions for which, at the given FAR, a search achieves 50% and 90% detection efficiencies. We present these conditions here in terms of ρ since it is the parameter that quantifies the detectability of a signal in optimum filtering. In tables 1 and 2, we present these ρ for a FAR $=$ 0.01 Hz. If a temporal coincidence between Virgo and one LIGO interferometers could be
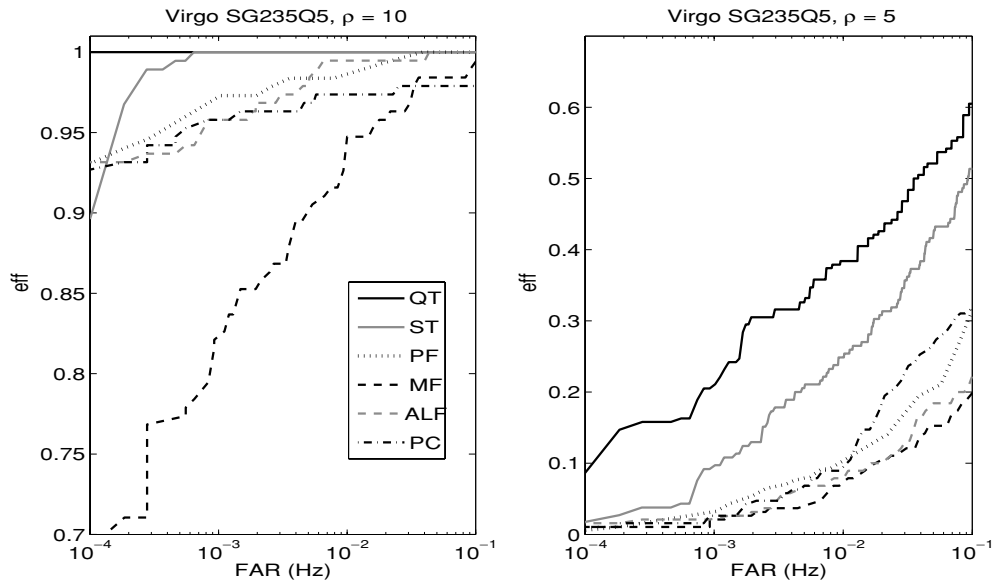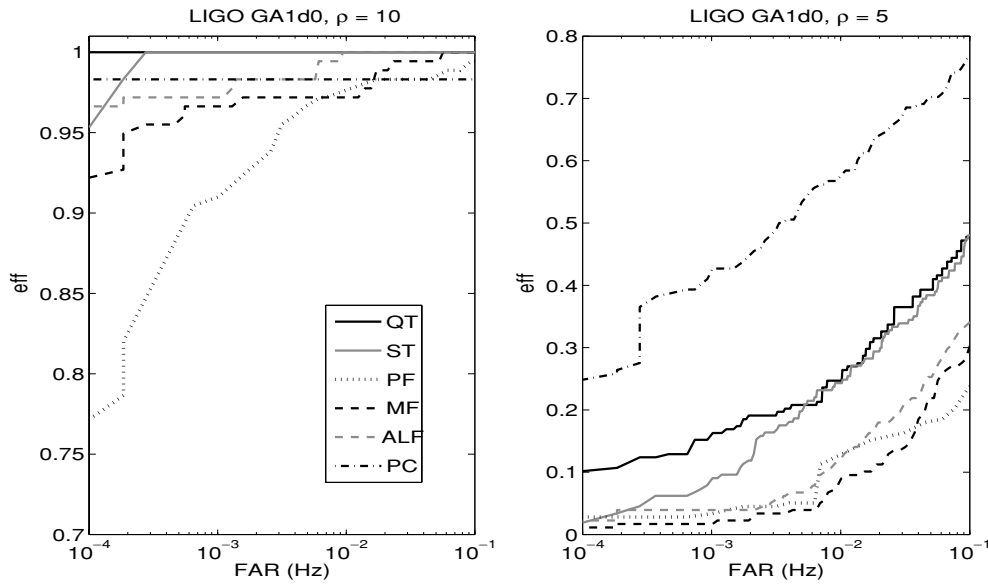
Figure 1. Virgo ROCs for GA235q5 and $\rho = 10$ (left) and 5 (right). MF is in black dashed, ALF in grey dashed, PC in black dashed–dotted, PF in black dotted, ST in solid grey and QT in solid black.
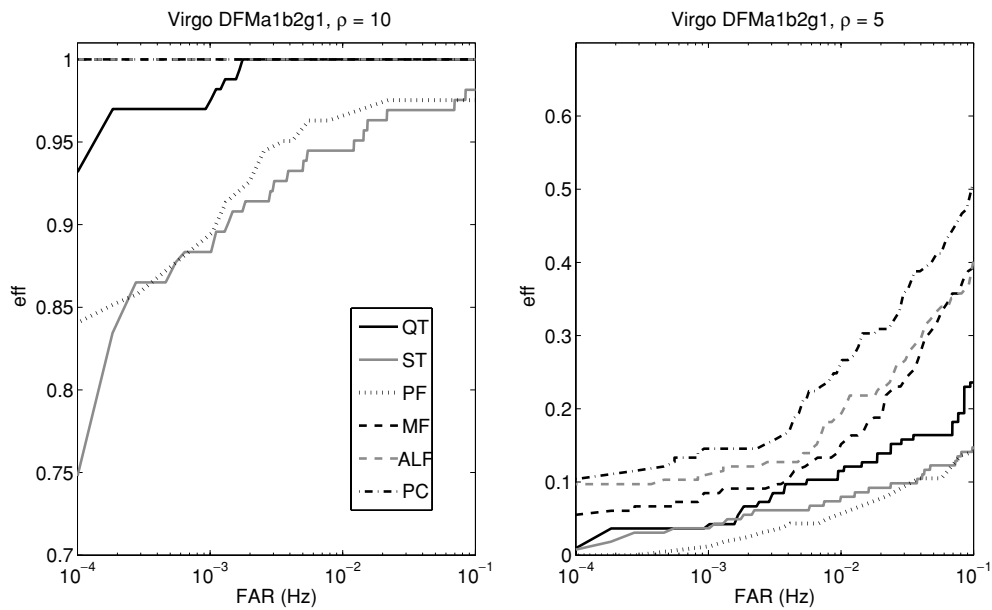

Figure 2. Virgo ROCs for DFMa1b2g1 and $\rho = 10$ (left) and 5 (right). MF is in black dashed, ALF in grey dashed, PC in black dashed–dotted, PF in black dotted, ST in solid grey and QT in solid black.

imposed the FAR above would correspond to less than a false event in 3 h. This is reasonable because if the noise triggers are generated randomly with a Poisson distribution, as is natural to expect for a stationary interferometer, the rate of coincidence can be estimated. The result is

Figure 3. LIGO ROCs for a 1 ms GA signal and $\rho = 10$ (left) and 5 (right). MF is in black dashed, ALF in grey dashed, PC in black dashed–dotted, PF in black dotted, ST in solid grey and QT in solid black.
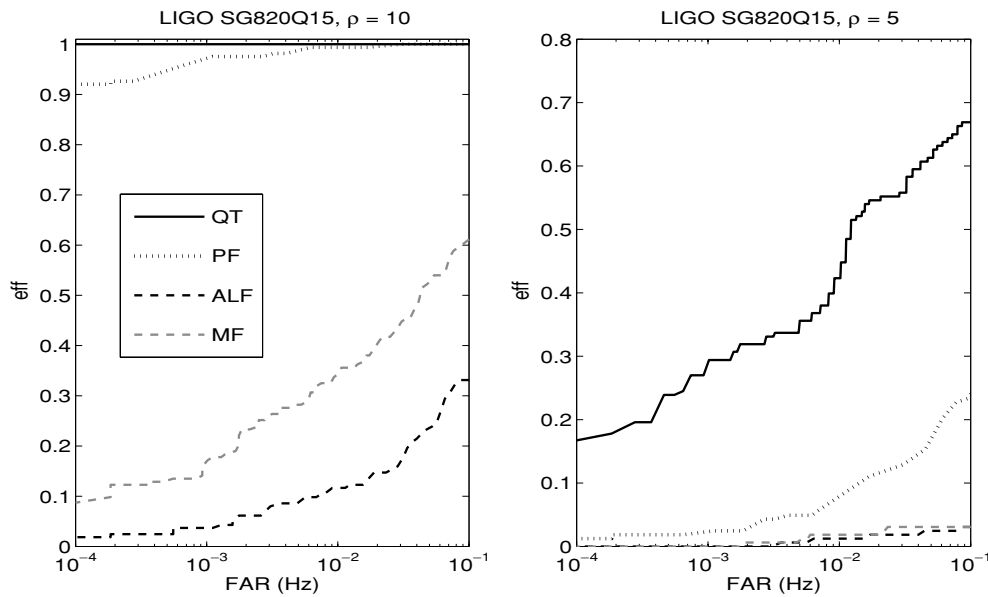
Figure 4. LIGO ROCs for a SG of central frequency equal to 829 Hz and $Q = 5$. MF in grey dashed, ALF in black dashed, PF in black dotted and QT in solid black.

roughly given by the product of the two rates by twice the length of the temporal coincidence window, which can be bounded with 50 ms for the searches involved in this study.

Explicitly, the values of the necessary ρ are computed in two steps: (a) the efficiencies for a given FAR and different ρ are extracted from the ROC curves and (b) a numerical

interpolation is performed with a least-square fit that minimizes the free parameters α, β and γ of the asymmetric sigmoid

$$E(\rho) = \frac{1}{1 + \left(\frac{\rho}{\gamma}\right)^{\alpha(1+\beta\tanh(\frac{\rho}{\gamma}))}}. \tag{2}$$

where γ corresponds to the value of ρ for which the detection efficiency $E = 0.5\beta$ is the parameter that describes the asymmetry of the sigmoid (and takes values between -1 and $+1$) and α describes the slope. The sigmoid function has been chosen because in past efficiency estimations using real data [11] it appeared to fit accurately the data and to extrapolate correctly within a few per cent the 50% efficiencies.

It is interesting to note from tables 1 and 2 that ST and QT, which are both based on a multiresolution decomposition of the time–frequency plane, have similar necessary ρ to obtain the 50% and 90% efficiencies. Time domain methods, similarly to the ROC curves, require higher ρ than for SG that is for GA and DFM waveforms.

The accuracy of the estimate of the arrival time of a pulse determines how strictly time consistency cuts can typically be involved in the postprocessing of the triggers and the angular resolution of a network of detectors. Explicitly, in order to identify the direction of arrival of a gravitational wave, the error in the peak time estimation needs to be much smaller than the travel times between LIGO interferometers ($\sim$10 ms) and between Virgo and one of the LIGO interferometers ($\sim$20 ms). A more quantitative analysis of the angular resolution is the goal of an ongoing project of this collaboration.

We present this information here in terms of the standard deviation and bias, both in milliseconds,

$$b = \frac{1}{N} \sum_{i=1}^{N} (\widehat{t_i} - t_i), \qquad \text{std} = \sqrt{\frac{1}{N-1} \sum_{i=1}^{N} ((\widehat{t_i} - t_i) - b)^2} \tag{3}$$

of the estimate of the peak arrival time $\widehat{t_i}$ with respect to the peak time of the injections t_i for $\rho = 10$ and FAR $= 0.1$ Hz (which provides the largest statistics among the FARs we studied). The accuracy of the searches appears to be similar between LIGO and Virgo data. Most of the methods can be used for a directional search on most of the waveforms since the time accuracy is typically smaller than the travel time between the interferometers. It is also worth noticing that the feasibility of directional searches rapidly degrades with decreasing ρ and that, if a method shows a systematic bias, this can usually be easily corrected.

For example, the PF computes the difference between the injected time and the arrival time. Since the adopted version of PF did not take into account the delay of the whitening filter, it is natural to expect a negative bias.

The current version of the algorithm has a better temporal resolution and takes into account the filter delay.

4. Conclusions

The main motivation of this work was to prepare the ground for joint searches of gravitational burst signals between the LIGO and Virgo interferometers. In particular, we built a joint framework where we learned how to exchange interferometer data, trigger files and analyse each other's data. We also gained a deeper understanding of the operational properties of our burst searches. This analysis showed that the performance of time domain and frequency domain methods have different strengths on different kinds of waveforms. The performance of the searches was fairly stable across both LIGO and Virgo data. The study of the arrival

time estimation accuracy sets the coincidence window for future multi-IFO studies and shows that most of the methods can be employed in directional searches given sufficient ρ. If a combination of the methods had to be applied on real data it would need to be studied as a search in itself. In particular, more extensive studies would be necessary to verify if the ROC curve of the combination of methods stands above the ROCs of the single searches. The next steps of this work are to extend the comparisons to directional searches performed on simulated data corresponding to a network of three interferometers, investigate the usefulness of combining methods and explore the use of consistency criteria other than time coincidence. It is also planned to expand the number of searches involved in the study.

Acknowledgments

References

[1] 2004 Proposal for Joint LIGO–Virgo Data Analysis, LIGO-T040137-08-Z and VIRGO-PLA-DIR-1000-201
[2] Beauville F *et al* 2005 *Class. Quantum Grav.* **22** S1149
[3] Abbot B *et al* (LIGO Scientific Collaboration) 2004 *Phys. Rev.* D **69** 122001
[4] Dimmelmeier H, Font J and Mueller E 2002 *Astron. Astrophys.* **393** 523–42
[5] Chatterji S, Blackburn L, Martin G and Katsavounidis E *Class. Quantum Grav.* **21** 2004
[6] Clapson A C *et al* 2005 *Class. Quantum Grav.* **22** S1381
[7] Guidi G M *et al* 2004 *Class. Quantum Grav.* **21** 815
[8] Arnaud N *et al* 1999 *Phys. Rev.* D **59** 082002
[9] Pradier T *et al* 2001 *Phys. Rev.* D **63** 042002
[10] Arnaud N *et al* 2003 *Phys. Rev.* D **67** 062004
[11] Abbot B *et al* (LIGO Scientific Collaboration) 2005 *Phys. Rev.* D at press (*Preprint* gr-qc/0505041)