Hierarchical Hough all-sky search for periodic gravitational waves in LIGO S5 data

Llucia Sancho de la Jordana for the LIGO Scientific Collaboration and the Virgo Collaboration

Departament de Física, Universitat de les Illes Balears, Cra. Valldemossa Km. 7.5, E-07122 Palma de Mallorca, Spain

E-mail: `llucia.sancho@uib.es`

Abstract. We describe a new pipeline used to analyze the data from the fifth science run (S5) of the LIGO detectors to search for continuous gravitational waves from isolated spinning neutron stars. The method employed is based on the Hough transform, which is a semi-coherent, computationally efficient, and robust pattern recognition technique. The Hough transform is used to find signals in the time-frequency plane of the data whose frequency evolution fits the pattern produced by the Doppler shift imposed on the signal by the Earth's motion and the pulsar's spin-down during the observation period. The main differences with respect to previous Hough all-sky searches are described. These differences include the use of a two-step hierarchical Hough search, analysis of coincidences among the candidates produced in the first and second year of S5, and veto strategies based on a χ^2 test.

1. Introduction

Spinning neutron stars are the most promising sources of continuous gravitational wave signals for ground-based interferometers such as GEO600 [1], LIGO [2] and VIRGO [3]. Using data from the different science runs, there have been two kinds of searches for gravitational waves from pulsars: i) targeted searches [4, 5, 6, 7, 8] for periodic gravitational radiation from pulsars whose parameters (sky-position and frequency evolution) are known through radio observations, and ii) searches for pulsars yet unobserved by radio telescopes [9, 10, 11, 12, 13, 14]. For this second kind of search the optimal method based on a coherent integration over the full observation time is computationally prohibitive for a wide-parameter search. Therefore all-sky searches require the use of semi-coherent techniques [15, 16, 17, 18, 19, 11], that are less sensitive for the same observation time but are computationally inexpensive, or hierarchical approaches that combine both methods, optimizing the sensitivity of a search for a given computational power [20, 19, 21, 22, 13]. Some of these semi-coherent methods operate on successive Short Fourier Transforms (SFTs) of the measured strain [23] searching for cumulative excess power from hypothetical periodic gravitational wave signal taking into account the Doppler modulation of the detected frequency due to the Earth's rotational and orbital motion with respect to the Solar System Barycenter, and the time derivative of the frequency intrinsic to the source. The Hough transform [10, 11, 15, 16] is an example of such a method.

Two flavors of the Hough transform have been developed and employed for different searches. The 'standard Hough' [10, 15], computes the cumulative excess power as a sum of binary zeroes and ones, where a SFT contributes a one if and only if the power exceeds a normalized power threshold, and the 'weighted Hough' [11, 17] in which the contribution of each SFT is weighted according to the noise and detector antenna pattern to maximize the signal-to-noise ratio. Using the 'weighted Hough' has two notable benefits: i) the gain in sensitivity, and ii) that this method can be used to analyze data from multiple detectors, since the different weights automatically take into account the different sensitivities.

This paper describes the pipeline employed to analyze the data from the fifth science run of the LIGO detectors to search for periodic gravitational waves using a hierarchical Hough transform, emphasizing the changes with respect to previous Hough searches [10, 11]. The paper is organized as follows: In section 2 we briefly summarize the basic principles of the Hough transform and its statistics. In section 3 we give an overview of the search pipeline and comment on the data employed and the parameter space of the search. In section 4 we describe two main features of our two-step hierarchical search, and we conclude in section 5.

2. The Hough transform

The starting point for the Hough transform is a set of N SFTs, short stretches of Fourier transformed data which are digitized by setting a threshold ρ_{th} on the normalized power ρ_k in the different frequency bins:

$$n_k^{(i)} = \begin{cases} 1 & \text{if} \quad \rho_k^{(i)} \geq \rho_{th} \\ 0 & \text{if} \quad \rho_k^{(i)} < \rho_{th} \end{cases} , \tag{1}$$

where

$$\rho_k = \frac{2|\tilde{x_k}|^2}{T_{\rm coh} S_n(f_k)} . \tag{2}$$

Here $\tilde{x_k}$ is the value of the Fourier transform in the k^{th} frequency bin corresponding to a frequency f_k, $T_{\rm coh}$ is the time baseline of the SFT (in our case $T_{\rm coh} = 1800$ s), and $S_n(f_k)$ is the single sided power spectral density of the detector noise at the frequency f_k.

For a given template, the Hough number count is a weighted sum of the binary zeroes and ones, along the corresponding time-frequency pattern

$$n = \sum_{i=0}^{N-1} w_k^{(i)} n_k^{(i)} , \tag{3}$$

where the weights are defined as

$$w_k^{(i)} \propto \frac{1}{S_k^{(i)}} \left\{ \left(F_{+1/2}^{(i)} \right)^2 + \left(F_{\times 1/2}^{(i)} \right)^2 \right\} , \tag{4}$$

being $F_{+1/2}^{(i)}$ and $F_{\times 1/2}^{(i)}$ the values of the beam pattern functions at the mid point of the i^{th} SFT. Thus, we take a binary count $n_k^{(i)}$ to have greater weight if SFT i has a lower noise floor or if, in the time-interval corresponding to this SFT, the beam pattern functions are larger for a particular point in the sky.

The weight normalization is chosen according to

$$\sum_{i=0}^{N-1} w_k^{(i)} = N . \tag{5}$$

With this normalization the Hough number count n lies within the range $[0, N]$. Note that the sensitivity of the search is governed by the ratios of the different weights, not by the choice of overall scale. The robustness of the Hough transform method in the presence of strong transient disturbances is not compromised by using weights because each SFT contributes at most w_i (which is of order unity) to the final number count.

The natural detection statistic is not the Hough number count n, but the *significance* of a number count, defined by:

$$s = \frac{n - \bar{n}}{\sigma}, \tag{6}$$

where $\bar{n}$ and σ are the expected mean and standard deviation for pure noise. Values of s can be compared directly across different templates characterized by different weight distributions and σ values. Furthermore, in the case of Gaussian noise, the relation between the significance and the false alarm probability α is given by:

$$\alpha = 0.5\,\mathrm{erfc}(s/\sqrt{2})\,. \tag{7}$$

Setting a threshold on the significance would then identify interesting candidates. We refer the reader to [11] for further details.

3. Description of the pipeline

This paper presents a new pipeline used to analyze the data from the LIGO detectors to search for continuous gravitational waves from isolated spinning neutron stars. The LIGO detector network consists of two interferometers in Hanford Washington, one 4-km and another 2-km (H1 and H2) and a 4-km interferometer in Livingston Louisiana (L1). The search described here is currently carried out over the entire sky using the data produced during LIGO's fifth science run (S5) that started on November 4, 2005 and ended on October 1, 2007, at initial LIGO's design sensitivity. Data from each of the three LIGO interferometers is used to perform the search.

The starting point is a collection of SFTs generated directly from the calibrated data stream, using 30-minute intervals of data for which the interferometer is operating in what is known as science mode. With this requirement, we search 32295 SFTs from the first year of S5 (11402 from H1, 12195 from H2 and 8698 from L1) and 35401 SFTs from the second year (12590 from H1, 12178 from H2 and 10633 from L1).

The search is performed in the frequency range $50-1000$ Hz and with the frequency's time derivative in the range -8.9×10^{-10} Hz s^{-1} to zero, being those values limited by the computational cost of the search. We use a uniform grid spacing equal to the size of a SFT frequency bin, $\delta f = 1/T_{\mathrm{coh}} = 5.556 \times 10^{-4}$ Hz. The resolution $\delta \dot{f}$ is given by the smallest value of $\dot{f}$ for which the intrinsic signal frequency does not drift by more than a frequency bin during the total observation time T_{obs}: $\delta \dot{f} = \delta f / T_{\mathrm{obs}} \sim 1.8 \times 10^{-11}$ Hz s^{-1}. This yields 51 spin-down values for each frequency. $\delta \dot{f}$ is fixed to the same value for the search on the first and the second year of S5 data, being T_{obs} the value for the first year. The sky resolution, $\delta\theta$, is frequency dependent, with the number of templates increasing with frequency, as given by Eq.(4.14) of Ref. [15]. This yields a resolution of about 9.3×10^{-3} rad at 300 Hz, which corresponds to $\sim 1.5 \times 10^{5}$ sky locations for the whole sky at that frequency.

The key difference from previous searches is that, starting from 30 min SFTs, we perform a multi-interferometer search analyzing separately the two years of the S5 run, and we study coincidences among the source candidates produced by the first and second years of data. This is inspired by a similar coincidence search using VIRGO data [24]. Furthermore, we use a χ^2 test adapted to the Hough transform searches, as described in [25], to veto potential candidates.

The approach used to analyze each year of data is based on a two-step hierarchical search for continuous signals from isolated neutron stars as described below. In both steps, the weighted Hough transform is used to find signals whose frequency evolution fits the pattern produced by the Doppler shift and the spin-down in the time-frequency plane of the data. The search is done by splitting the frequency range in 0.25 Hz bands and using the SFTs from multiple interferometers.

In the first stage, we break up the sky into smaller patches with frequency dependent size. The size of the sky-patches ranges from ~ 0.4 rad $\times\, 0.4$ rad at 50 Hz to ~ 0.07 rad $\times\, 0.07$ rad at 1 kHz. Ideally, to obtain the maximum increase in sensitivity, we should calculate the weights, based both on the noise and the antenna pattern, for each sky-location. In practice, we calculate the weights just once for the center of each sky-patch. In this first stage, we perform the Hough multi-interferometer search using the traditional *look up table* approach, that enormously reduces the computational cost (see [15] for a detailed description). But limitations on the memory of the machines constrain the volume of data (i.e., the number of SFTs) that can be analyzed at once and the parameter space (e.g., size and resolution of the sky-patches and number of spin-down values) we can search over. For this reason, in this first stage, the Hough transform is not applied using all the available SFTs, but selecting the best 15000 SFTs for each frequency band and sky-patch. A top-list keeping the best 1000 candidates is produced for each 0.25 Hz band.

In a second stage, we recompute the Hough significance of each candidate in the top-list using the complete set of available SFTs from all the interferometers, and at the same time we reduce the mismatch of the template, i.e., we calculate the number count and the corresponding significance without the roundings introduced by the *look up table* approach and by recomputing the weights for the precise sky location and not for the center of the corresponding patch, as previously done. For each candidate, we also compute the χ^2 value that will help in vetoing the resulting candidates.

4. A two-step hierarchical Hough search

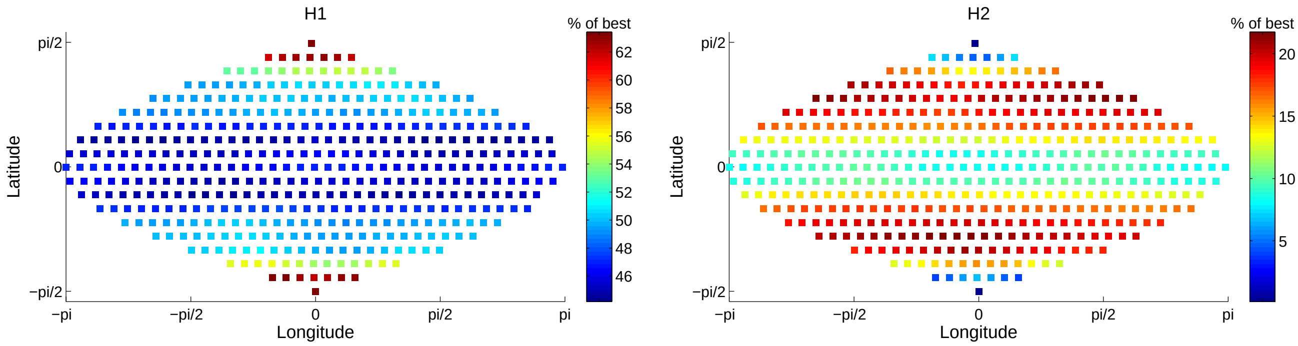
In this section we describe in more detail the two main features of the two-step hierarchical Hough search, i.e., the selection of the best SFT data and the comparison of the significance values produced in both steps.

4.1. Selection of the best SFT data

The first of the two main changes to the Hough code consists in selecting the best SFTs based on the weights given by Eq. (4). In this way, when doing the all-sky search for each 0.25 Hz band, for each sky-patch, we will keep the 15000 SFTs that have lower noise and that are more sensitive at that sky location.

Figure 1 shows the percentage of SFTs from each of the three LIGO detectors (H1, H2 and L1) that have been selected when doing the Hough search on a 0.25 Hz band at a frequency of 420 Hz for the S5 first year of data. At this frequency, the number of sky-patches is 426 with a size of ~ 0.17 rad $\times\, 0.17$ rad. The figure shows the percentage of SFTs from each detector on each sky-patch. In this figure we can see that, for this particular frequency, the detector that contributes the most at almost all the sky locations is H1. The maximum contribution of H1, about 64%, is at the poles, L1 has it maximum contribution, about 46%, around the equator and H2 contributes at most 23% of the SFTs in the region in between.

The selection of SFTs has been done based on the total weights, that depend both on noise floor and antenna pattern. If this selection had been based only upon the weights due to the noise floor, the H2 detector would not have contributed at all in this first stage. By having the weights antenna pattern dependent, H2 has a certain contribution in some sky locations where it is more sensitive.



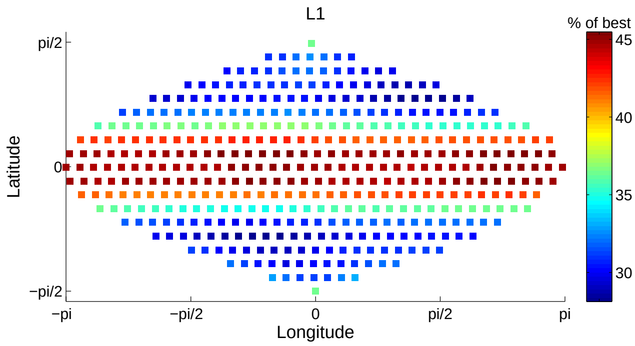


Figure 1. Percentage of SFTs that each detector has contributed at each particular sky-patch when selecting the best 15000 SFTs based on the Hough weights. These figures correspond to an all-sky search performed on a 0.25 Hz band at 420 Hz for the S5 first year of data.

4.2. Comparison of the significance values

In the first stage of the hierarchical search, the Hough significance is computed using the *look up table* approach to enormously reduce the computational cost. The code uses the selected 15000 best SFTs to compute the Hough significance and make a list of interesting candidates.

In the second stage of the search, for each candidate stored in the top-list produced for each 0.25 Hz band, the code recomputes the Hough significance without any rounding, and using all the SFTs from the three LIGO detectors. In this stage, the significance is computed directly from the number count that is obtained from the digitized time-frequency plane by summing weighted binary zeros or ones along the expected path of the frequency of a hypothetical periodic gravitational wave signal, taking into account the Doppler shift and the spin-down.

Figure 2 shows how this recomputation affects the previous value of the significance in the case where an injected pulsar was present and in the case of a frequency band free of spectral disturbances. To illustrate the effect, we have used a 'quiet band' at 205.25 Hz and a band with an injected pulsar at 575 Hz. In both bands we perform an all-sky search and compare the significances computed on the first and second stages. The 'old significance' corresponds to the one computed in the first stage, using the *look up table* approach and using only the best 15000 SFTs. The 'new significance' corresponds to the one computed in the second stage using all the SFTs and without any rounding. In the same figure we also provide the histograms of the differences between the 'old' and 'new' significances, showing that in the case where we have the hardware injected signal, the values of the 'new' significance are, in most cases, higher than the 'old' values. While in the case of the 'quiet' band, the 'new significance' becomes smaller than the 'old' one in the majority of the cases.

5. Conclusions and future work

In this paper, we have presented the improved Hough search pipeline, which is being used to analyze data from the fifth science run of the LIGO interferometers, describing in detail the two main features of the two-step hierarchical search, i.e., the selection of SFT data and

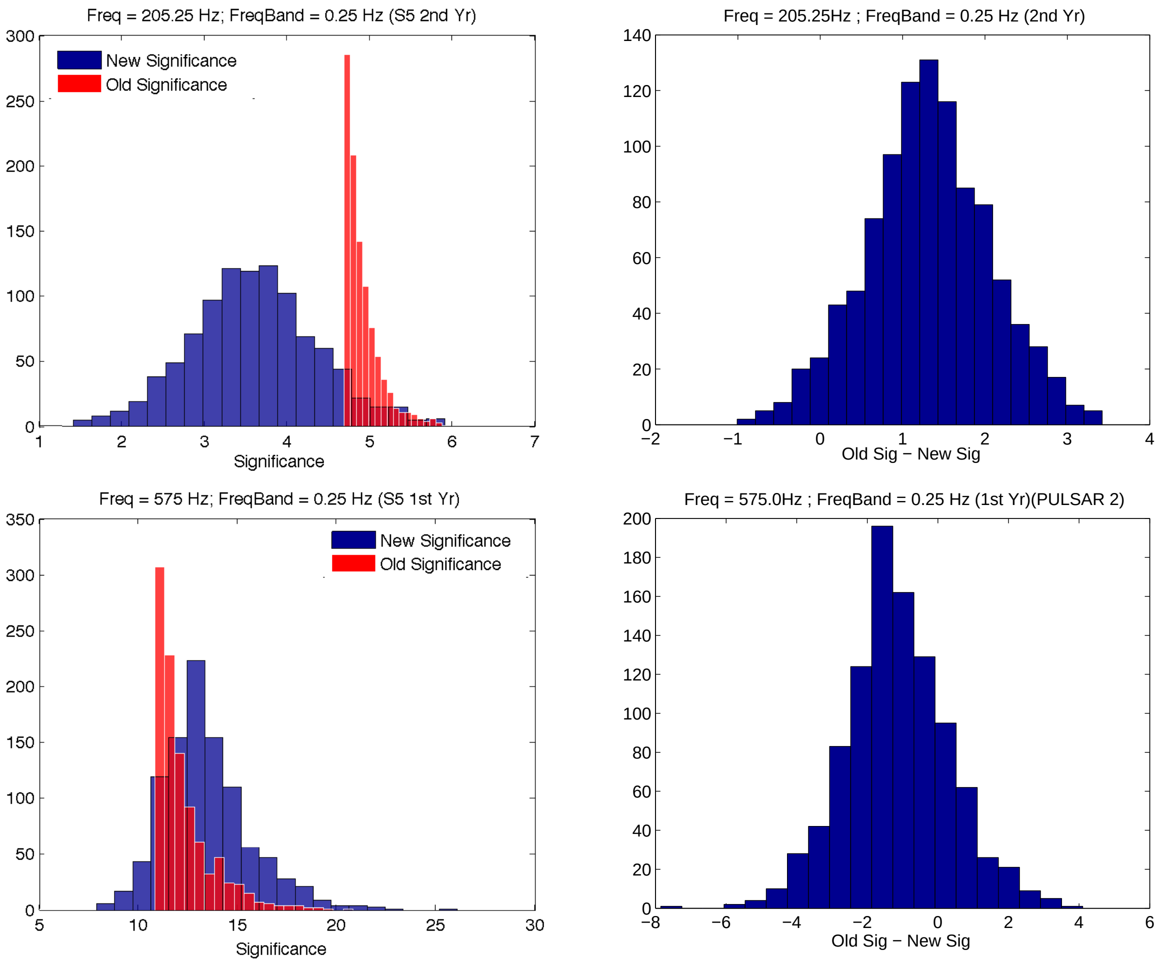


Figure 2. Comparison of the distribution of the significance values when calculated in the first stage using the *look up table* approach ('old significance') and after recomputing its value using all the SFTs and without any rounding ('new significance'). These results correspond to an all-sky search on a 0.25 Hz band at a frequency of 205.25 Hz for the S5 2nd year of data (top panels), where no relevant spectral disturbances were present, and at 575 Hz for the 1st year of data (bottom panels), containing a hardware injected pulsar.

recomputation of the significance values. We have shown how these improvements perform on the data, showing examples at some particular frequencies, either with interesting artifacts, such as hardware injected pulsars, or others where we expect approximately Gaussian noise. Work is in progress to analyze the coincidences among the candidates produced in the first and second year of S5 for the full parameter space, and to compute astrophysical upper limits using the search pipeline presented in this paper.

Acknowledgments

The authors gratefully acknowledge the support of the United States National Science Foundation for the construction and operation of the LIGO Laboratory, the Science and Technology Facilities Council of the United Kingdom, the Max-Planck-Society, and the State of Niedersachsen/Germany for support of the construction and operation of the GEO600 detector, and the Italian Istituto Nazionale di Fisica Nucleare and the French Centre National de la Recherche Scientique for the construction and operation of the Virgo detector. The authors also gratefully acknowledge the support of the research by these agencies and by the Australian Research Council, the Council of Scientic and Industrial Research of India, the Istituto Nazionale di Fisica Nucleare of Italy, the Spanish Ministerio de Educacion y Ciencia, the Conselleria dEconomia Hisenda i Innovacio of the Govern de les Illes Balears, the Foundation for Fundamental Research on Matter supported by the Netherlands Organisation for Scientic Research, the Royal Society, the Scottish Funding Council, the Polish Ministry of Science and Higher Education, the FOCUS Programme of Foundation for Polish Science, the Scottish Universities Physics Alliance, The National Aeronautics and Space Administration, the Carnegie Trust, the Leverhulme Trust, the David and Lucile Packard Foundation, the Research Corporation, and the Alfred P. Sloan Foundation. LIGO Document No. ligo-p0900212

6. References

[1] Luck H *et al.* 2006 *Class. Quant. Grav.* **23** S71–S78
[2] Abbott B *et al.* (The LIGO Scientific Collaboration) 2009 *Rept. Prog. Phys.* **72** 076901 (*Preprint* `0711.3041`)
[3] Acernese F *et al.* 2008 *Class. Quant. Grav.* **25** 114045
[4] Abbott B *et al.* (The LIGO Scientific Collaboration) 2004 *Phys. Rev.* **D69** 082004 (*Preprint* `gr-qc/0308050`)
[5] Abbott B *et al.* (The LIGO Scientific Collaboration) 2005 *Phys. Rev. Lett.* **94** 181103 (*Preprint* `gr-qc/0410007`)
[6] Abbott B *et al.* (The LIGO Scientific Collaboration) 2007 *Phys. Rev.* **D76** 042001 (*Preprint* `gr-qc/0702039`)
[7] Abbott B *et al.* (The LIGO Scientific Collaboration) 2008 *Astrophys. J.* **683** L45–L50 (*Preprint* `0805.4758`)
[8] Abbott B *et al.* (The LIGO Scientific Collaboration and The Virgo Collaboration) 2009 (*Preprint* `0909.3583`)
[9] Abbott B *et al.* (The LIGO Scientific Collaboration) 2007 *Phys. Rev.* **D76** 082001 (*Preprint* `gr-qc/0605028`)
[10] Abbott B *et al.* (The LIGO Scientific Collaboration) 2005 *Phys. Rev.* **D72** 102004 (*Preprint* `gr-qc/0508065`)
[11] Abbott B *et al.* (The LIGO Scientific Collaboration) 2008 *Phys. Rev.* **D77** 022001 (*Preprint* `0708.3818`)
[12] Abbott B *et al.* (The LIGO Scientific Collaboration) 2009 *Phys. Rev.* **D79** 022001 (*Preprint* `0804.1747`)
[13] Abbott B P *et al.* (The LIGO Scientific Collaboration) 2009 *Phys. Rev.* **D80** 042003 (*Preprint* `0905.1705`)
[14] Abbott B *et al.* (The LIGO Scientific Collaboration) 2009 *Phys. Rev. Lett.* **102** 111102 (*Preprint* `0810.0283`)
[15] Krishnan B *et al.* 2004 *Phys. Rev.* **D70** 082001 (*Preprint* `gr-qc/0407001`)
[16] Palomba C, Astone P and Frasca S 2005 *Class. Quant. Grav.* **22** S1255–S1264
[17] Sintes A M and Krishnan B 2006 *J. Phys. Conf. Ser.* **32** 206–211 (*Preprint* `gr-qc/0601081`)
[18] Antonucci F *et al.* 2008 *Class. Quant. Grav.* **25** 184015 (*Preprint* `0807.5065`)
[19] Brady P R and Creighton T 2000 *Phys. Rev.* **D61** 082001 (*Preprint* `gr-qc/9812014`)
[20] Papa M A, Schutz B F and Sintes A M 2000 *ICTP Lecture Notes Series Vol. III, edited by V. Ferrari, J. C. Miller and L. Rezzolla (Italy)* 431–442 (*Preprint* `gr-qc/0011034`)
[21] Frasca S, Astone P and Palomba C 2005 *Class. Quant. Grav.* **22** S1013–S1019
[22] Cutler C, Gholami I and Krishnan B 2005 *Phys. Rev.* **D72** 042004 (*Preprint* `gr-qc/0505082`)
[23] Astone P, Frasca S and Palomba C 2005 *Class. Quant. Grav.* **22** S1197–S1210
[24] Acernese F *et al.* 2007 *Class. Quant. Grav.* **24** S491–S499
[25] Sancho de la Jordana L and Sintes A M 2008 *Class. Quant. Grav.* **25** 184014 (*Preprint* `0804.1007`)