LIGO-P070022-00-Z

Search algorithm for the detection of long-duration narrow-band transients in GW interferometers

L. Matone, S. Márka

Columbia University, New York, NY 10027, USA

E-mail: matone@phys.columbia.edu

Abstract. The recent observation of Quasi-Periodic Oscillations (QPOs) in the X-ray light curve of Soft Gamma-ray Repeaters (SGRs), which fall within LIGO's sensitivity band, prompted us to search for Gravitational Waves (GWs) associated with them. We describe the corresponding search algorithm that is based on the energy measurement in a single detector signal stream. We analytically derive the search sensitivity and compare it to the numerical sensitivity provided by the analysis pipeline with simulated detector noise. We found excellent agreement between the two approaches reassuring us that the analysis method is well-understood and implemented properly. The detector noise is approximated as white Gaussian with a strain equivalent spectral noise density of 10^{-22} strain $\mathrm{Hz}^{-1/2}$, similar to the noise floor of the H1 LIGO detector at the time of the SGR $1806-20$ hyperflare event of 27 December 2004. As a trial model we use a hypothetical 100 Hz QPO lasting for 50 s. The corresponding energy sensitivity is found to be $E^{sens} = 3.63 \times 10^{-43}$ strain^2 Hz^{-1} which, in terms of amplitude is $h^{sens}_{rss-det} = 6.06 \times 10^{-22}$ strain $\mathrm{Hz}^{-1/2}$. We show that, besides being simple and flexible, the algorithm is sensitive to a wide range of waveforms obeying the time and bandwidth requirements.

1. Introduction

Soft Gamma-ray Repeaters (SGRs), one class of target sources this algorithm addresses, are objects that emit short-duration X and gamma-ray bursts at irregular intervals (see [1] for a review). At times these sources emit giant flares lasting hundreds of seconds (see for example [2–4]) with peak electromagnetic luminosities reaching 10^{47} erg/s [5].

On 27 December 2004, SGR $1806-20$ emitted a record flare significantly more luminous than any previous transient event observed in our Galaxy. In the context of the *magnetar* model [6] the object is a neutron star with a high magnetic field ($B \sim 10^{15}$ G). In this model the giant flares are generated by the catastrophic rearrangement of the neutron star's crust and magnetic field: a *starquake* [7; 8].

The Quasi-Periodic Oscillations (QPOs) in the pulsating tail of the 27 December 2004 event were observed by the *Rossi X-Ray Timing Explorer (RXTE)* and *Ramaty High Energy Solar Spectroscopic Imager (RHESSI)* satellites [9–11]. It has been suggested that the star's seismic modes, excited by this catastrophic event, might drive the observed QPOs [9–13] (more discussion on this topic can be found in Ref. [14; 15]), which lead us to develop a search algorithm for the detection of plausible narrow-band transients in GW interferometers associated with them.

In the absence of detailed theoretical models of GW emission from magnetar QPOs, we keep the GW search as broad and sensitive as possible, while following the QPO signatures observed in the electromagnetic spectrum both in frequency and time interval. On December 27, 2004, only the LIGO Hanford (WA) 4 km detector (H1) (see Ref. [16] for a description of the detector) was collecting low noise data thus restricting us to an *excess power* type search, variants of which are described in Ref. [17–19]. The described algorithm, used in the astrophysical results described in Ref. [20], can easily be adapted to use data streams from multiple interferometers.

2. Description of the search and sensitivity estimates

The algorithm analyzes a single data stream at multiple frequency bands by calculating the energy excess in the time domain. This is done by measuring the energy of the GW detector output for a given bandwidth, Δf, and time slice, Δt, corresponding to the observed QPO measurements in the electromagnetic spectrum [9–11]. We measure the energy for a given time interval at the observed QPO frequencies for a given bandwidth and compare it to the energy measured in adjacent frequency bands not related to the QPO. The energy excess is then calculated for each time-frequency volume of interest.

To estimate the search sensitivity, we first analytically derive the expected sensitivity by assuming white stationary Gaussian statistics for the noise floor of the detector (the search sensitivity using real LIGO data has been explored in Ref. [20]). To validate the analysis, the analytical expectations are compared to numerical results provided by the search pipeline by adding (or *injecting*) arbitrary simulated gravitational waveforms to simulated detector noise.

2.1. Energy statistics using a single frequency band [17–19]

Let X_i be a normally distributed and independent random variable with mean μ_i and variance σ_i^2. The *normalized* sum of squares X^2, defined as

$$X^2 = \sum_{i=1}^{\nu} \left(\frac{X_i}{\sigma_i} \right)^2 \tag{1}$$

which is also a random variable distributed according to a non-central χ^2_ν distribution of ν degrees of freedom and non-central parameter λ [21] where

$$\lambda = \sum_{i=1}^{\nu} \left(\frac{\mu_i}{\sigma_i} \right)^2 \tag{2}$$

The mean μ_{X^2} and variance $\sigma^2_{X^2}$ of the resulting distribution is

$$\begin{aligned} \mu_{X^2} &= \nu + \lambda \\ \sigma^2_{X^2} &= 2(\nu + 2\lambda) \end{aligned} \tag{3}$$

In the limit $\nu \to \infty$ the χ^2_ν distribution can be approximated by a Gaussian distribution.

If we assume that the GW output y_i is band-limited white noise, sampled at frequency F_s, of amplitude spectral density $\tilde{n} = \sigma\sqrt{2/F_s}$, and extracted from a normal distribution of mean μ and variance σ^2, then the signal energy is also a random variable distributed according to a non-central χ^2 distribution. This is seen by the functional form of the signal energy E, measured in the time interval Δt and defined as

$$E = \frac{1}{F_s} \sum_{i=1}^{N} y_i^2 \tag{4}$$

which is identical to the functional form of Eq.(1) where N is the number of points sampled in the time segment $\Delta t = N/F_s$.

Setting $\mu = 0$ (the GW output y_i has a null DC component) and using Eq.(3) and Eq.(2), the mean and variance of the energy distribution are

$$\begin{aligned} \mu_E &= \frac{N\ \sigma^2}{F_s} = \tilde{n}^2\ \Delta t\ \Delta f \\ \sigma_E^2 &= 2\ \frac{N\ \sigma^4}{F_s^2} = \tilde{n}^4\ \Delta t\ \Delta f \end{aligned} \tag{5}$$

where the bandwidth Δf is set to the Nyquist frequency $F_s/2$ and $\sigma^2 = \Delta f\ \tilde{n}^2$.

Eq.(5) statistically describes the expected background and determines the search sensitivity. In particular, this sensitivity will depend on the background variance which, in turn, depends on the fourth power of the detector's noise floor $\tilde{n}^4$ and on the time-frequency volume $\Delta t \times \Delta f$ used in the search.

This sensitivity can be quantified by injecting a waveform into the data stream of the detector. Referring to μ_E^b and $(\sigma_E^b)^2$ as the *background* (absence of a signal) mean and variance, the resulting energy distribution is shifted by the injected energy E^{inj}

$$\mu_E^{inj} = \mu_E^b + E^{inj} \tag{6}$$

as shown in Fig.(1).

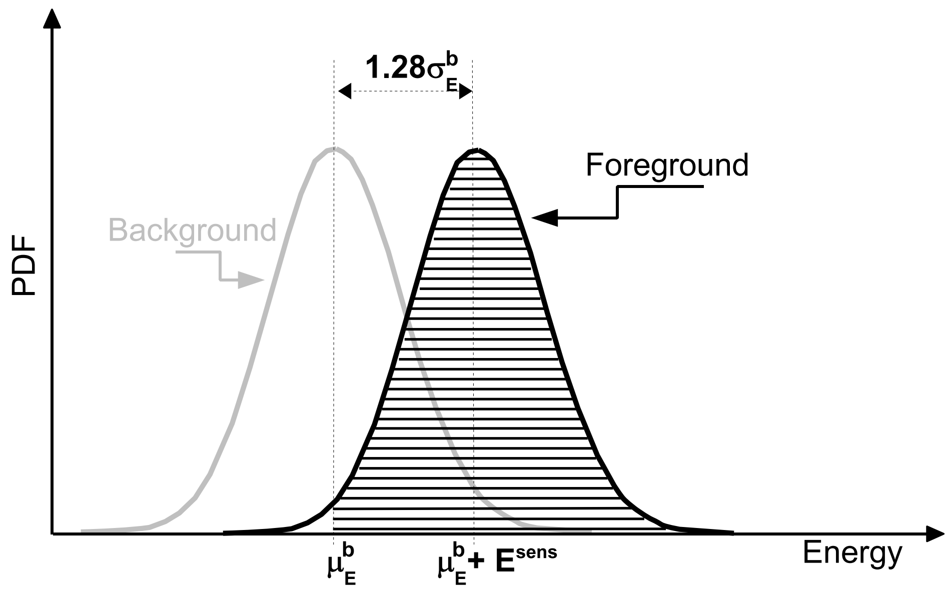

Figure 1. The energy distribution with and without injected waveforms. The diagram shows how the variance of the distribution determines the sensitivity of the search. In this work we define this sensitivity to be the injected energy such that 90% of the resulting signal is above the background median (Eq.(7))

We define the search sensitivity E^{sens} as the added energy E^{inj} that gives rise to a distribution with 90% of it being above the background median:

$$E^{sens} = 1.28\ \sigma_E^b \tag{7}$$

Using Eq.(5), this reduces to

$$E^{sens} = 1.28\ \tilde{n}^2 \left(\Delta t\ \Delta f\right)^{1/2} \tag{8}$$

This sensitivity, as previously discussed, is proportional to the square of the detector noise equivalent power spectral density $\tilde{n}^2$ and to the square root of the time-frequency volume in use. Fig.(1) provides a graphical description of this sensitivity.

2.2. Excess energy statistics using multiple frequency bands

Due to the narrow band nature of the search we can approximate the detector noise as white noise; i.e. there is no significant frequency dependence of the noise power over the search band. On the other hand, stationary independent Gaussian variables do not approximate well the detector's noise in the presence of both non-stationarity and occasional broad-band instrumental transients, that are expected to degrade the search sensitivity significantly, if not mitigated.

To minimize this effect the search algorithm measures the energy *difference* between the band of interest and the average of neighboring bands (not related to the observed QPO). As a consequence the sensitivity is slightly lower than the theoretical sensitivity, given by Eq. (8), under ideal conditions. However, in the presence of non-stationarity and broad-band noise contamination, the use of a multi-band energy measure outperforms the single one as seen in the use of the algorithm in real LIGO data in Ref. [20].

The use of multi-bands to measure an energy *excess* ΔE is defined as

$$\begin{aligned}\Delta E &= E - E_{avg} \\ E_{avg} &= \frac{1}{2}\Big(E_{low} + E_{up}\Big)\end{aligned} \tag{9}$$

where E is the energy in the band of interest, $E_{low,up}$ is the energy of the lower and upper sideband and E_{avg} is the energy average of the two adjacent bands.

Assuming that each of the three energy measures, E, E_{low} and E_{up}, are independent and extracted from the same parent population of mean μ_E and variance σ_E^2, then it can be shown that

$$\begin{aligned}\sigma^2_{E_{avg}} &= \frac{1}{4}\,\sigma^2_E + \frac{1}{4}\,\sigma^2_E = \frac{1}{2}\,\sigma^2_E \\ \sigma^2_{\Delta E} &= \sigma^2_E + \frac{1}{2}\,\sigma^2_E = \frac{3}{2}\,\sigma^2_E\end{aligned} \tag{10}$$

The resulting energy spread increases by a factor of $\sqrt{1.5}$ thus decreasing the energy sensitivity by the same amount. Substituting $\sigma_E^b = \sigma_{\Delta E}$ in Eq.(7), the multi-band energy sensitivity is

$$E_{sens} = 1.28\,\sqrt{\frac{3}{2}}\;\tilde{n}^2\,\Big(\Delta t\;\Delta f\Big)^{1/2} \tag{11}$$

This is the expected energy sensitivity for the search algorithm presented in this work.

3. Description of the pipeline

In this section we describe the search pipeline, written in Matlab and running in the time domain, for the measure of *excess* energy exploiting multiple frequency bands.

A block diagram of the analysis pipeline is shown in Fig.(2) where in this case simulated detector noise is used. The QPO observations in the electromagnetic spectrum determine the time-frequency volume of the search. To provide an estimate of the search sensitivity, an arbitrary simulated gravitational waveform can be injected to each data segment. The data stream, with and without injections, is then conditioned or band-pass filtered to select the three frequency bands of interest: the QPO band and the two adjacent frequency bands.

After the conditioning procedure is complete, the data stream is pushed through the search algorithm, which computes the power in each segment for the three frequency bands of interest and the excess power in the segment. The excess energy distribution for the background and foreground are then compared.

3.1. Data conditioning and search algorithm

The conditioning procedure consists of zero-phase filtering [22] of the data with three different band-pass Butterworth filters. The first band-pass filters the data around the

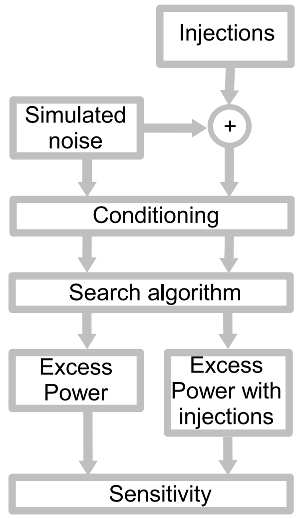

Figure 2. A block diagram of the analysis sketching the signal flow.

QPO frequency of interest with a predefined bandwidth. This bandwidth depends on the observed QPO width and on the fact that the QPOs have been observed to evolve in frequency. For this study, the bandwidth is set to 10 Hz (which is sufficiently over the measured QPO FWHM [11] and frequency evolution). The data is also filtered to select the two adjacent frequency bands with identical bandwidths to the chosen QPO band.

The algorithm takes the difference in energy between a band centered at a frequency $f_{\rm qpo}$ and the average of the two frequency bands adjacent to the QPO frequency band, also of bandwidth Δf, typically centered at $f_{\pm} = f_{\rm qpo} \pm \Delta f$.

After band-pass filtering, we are left with three channels for each QPO: $c_{\rm qpo}(t)$, $c_{+}(t)$, and $c_{-}(t)$. The energy for the QPO interval for each of these channels is

$$E_{{\rm qpo},\pm} = \int_{t_{\rm start}}^{t_{\rm end}} (c_{{\rm qpo},\pm})^2 dt \tag{12}$$

The energy excess, as defined in Eq. (9), is

$$\Delta E = E_{\rm qpo} - E_{\rm avg} \tag{13}$$

where $E_{\rm avg} = (E_{+} + E_{-})/2$ is the average of the adjacent bands. We refer to the resulting set of ΔE calculated as the *background* for no injections and *foreground* with injections.

In order to numerically estimate the sensitivity of the search and validate the pipeline, we injected Sine-Gaussian (SG) waveforms. They can be parameterized as

$$h_{\rm det}(t) = A\ \sin\left(2\pi f_{\rm c} t + \phi\right)\, e^{-(t-t_0)^2/\tau^2} \tag{14}$$

where A is the waveform peak amplitude, $f_{\rm c}$ is the waveform central frequency, $Q = \sqrt{2}\pi\tau f_{\rm c}$ is the quality factor, τ is the $1/e$ decay time, ϕ is an arbitrary phase and t_0 indicates the waveform peak time. Each waveform is added directly to the same simulated data segments before data conditioning is applied.

The strength of the injected strain (at the detector) $h_{\rm det}(t)$ is defined by its *root-sum-square* (rss) amplitude, or

$$h_{\rm rss-det} = \sqrt{\int_{t_1}^{t_1+\Delta t} |h_{\rm det}(t)|^2 {\rm dt}} \tag{15}$$

integrated over the interval Δt where t_1 indicates the start of a segment in the background region. The search sensitivity to a particular waveform, $h^{\rm sens}_{\rm rss-det}$, is defined as the injected amplitude $h_{\rm rss-det}$ such that 90% of the resulting $\Delta\mathcal{P}$ distribution is above the background median.

4. Monte Carlo validation

We describe the detector noise as white Gaussian strain equivalent noise of 10^{-22} strain Hz$^{-1/2}$. This floor is similar to the noise floor level of the 4km Hanford LIGO detector (H1) at the time of the hyperflare. We also assume a trial QPO model at 100 Hz and lasting for 50 s (similar to the observations in Ref. [9]). In this case, the band center frequencies, bandwidths and signal durations are set to $f_{\rm qpo} = 100$ Hz, $f_- = 90$ Hz, $f_+ = 110$ Hz, $\Delta f = 10$ Hz and $\Delta t = 50$ s.

There are effects that need to be taken into account for the comparison, one of them being being the filter shape or filter roll off. This roll off, if not aggressive enough, modifies the filter bandwidth and generates an effective bandwidth Δf_{eff}. This bandwidth can be quantified using the filter transfer function and comparing it to an ideal filter shape (with infinite rool off) as shown in Fig.(3). In particular, one defines this effective bandwidth as the power ratio of the unfiltered signal and the bandpassed one. Using Fig.(3), which shows the filter used for our trial QPO model, we calculate

$$\Delta f_{eff} = 10.20 \text{ Hz} \tag{16}$$

The simulation results showed a spread increase to the foreground distribution which arises from the injected waveform. This increase is quantified by taking the ratio of the foreground to the background standard deviation

$$\frac{\sigma_E^{inj}}{\sigma_E^{bg}} = 1.076 \tag{17}$$

Due to this increase, Eq. (10) is modified and the variance foreground increase is given

$$\begin{aligned}\sigma^2_{\Delta E} &= 1.076^2\ \sigma^2_E + \frac{1}{2}\ \sigma^2_E \\ &= 1.576\ \sigma^2_E\end{aligned} \tag{18}$$

The expected sensitivity E^{sens}, following Eq. (11), is

$$\begin{aligned}E^{sens} &= 1.28 \times \sqrt{1.576} \times \sqrt{50 \times 10.2} \times 10^{-44} \text{ strain}^2 \text{ Hz}^{-1} \\ &= 3.6289 \times 10^{-43} \text{ strain}^2 \text{ Hz}^{-1}\end{aligned} \tag{19}$$

The numerical result of the search under the same condition is shown in Fig.(4), together with one standard deviation statistical error uncertainty is

$$E_{sens} = 3.66 \times 10^{-43} \pm 3 \times 10^{-45} \text{strain}^2 \text{ Hz}^{-1} \tag{20}$$

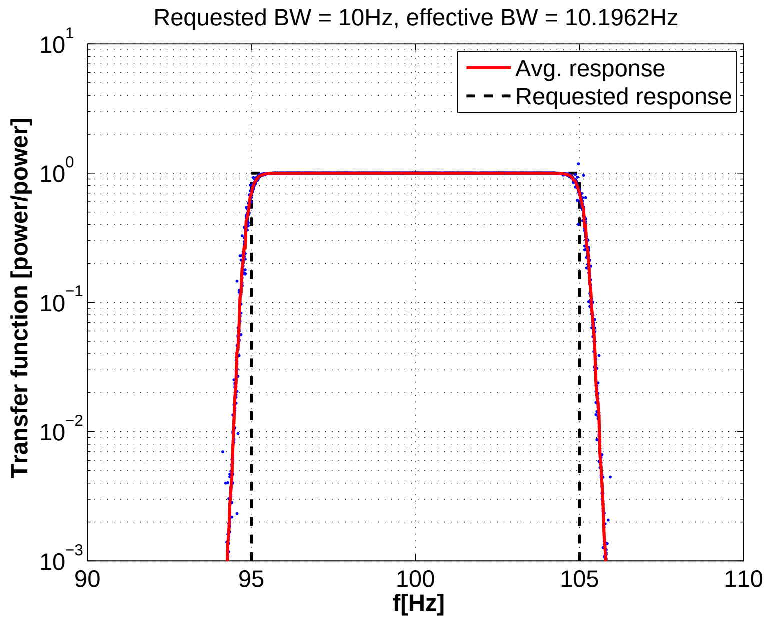

Figure 3. The transfer function of the band-pass filter used for the QPO band of interest, centered at 100 Hz with bandwidth of 10 Hz. The dashed line indicates an ideal bandpass filter. The black line shows the filter used for this study. By taking the area of the filter curve, an effective bandwidth can be defined and it is found to be 10.20 Hz.

The agreement between the expected sensitivity shown in Eq.(19) and the Monte Carlo simulation result of Eq.(20) is at the $\sim 1\%$ level and is found to be consistent to the statistical determination of the distribution parameters. In terms of h_{rss} (amplitude) sensitivity, the search algorithm presented here provides a sensitivity of

$$h^{sens}_{rss-det} = 6.06 \times 10^{-22} \text{ strain Hz}^{-1/2} \tag{21}$$

Astrophysical results were made with the use of the present search pipeline (see Ref. [20]). Fig.(3) of Ref. [20] plots the search sensitivity, using LIGO background data at the time of the hyperflare event of 27 December 2004, to different waveform families. The result shown indicates an approximately constant search sensitivity over all of the waveforms considered and in excellent agreement with the numerical and analytical estimates presented in this work.

5. Conclusion

In this work we presented the analytical and numerical (via the use of a Monte Carlo) sensitivity of a GW search, based on an *excess power* type search, to the observed electromagnetic QPOs in the SGR $1806-20$ hyperflare event of 27 December 2004.

We analytically derive the expected sensitivity and describe the sensitivity of our pipeline via Monte Carlo simulation. The agreement between the analytical and Monte

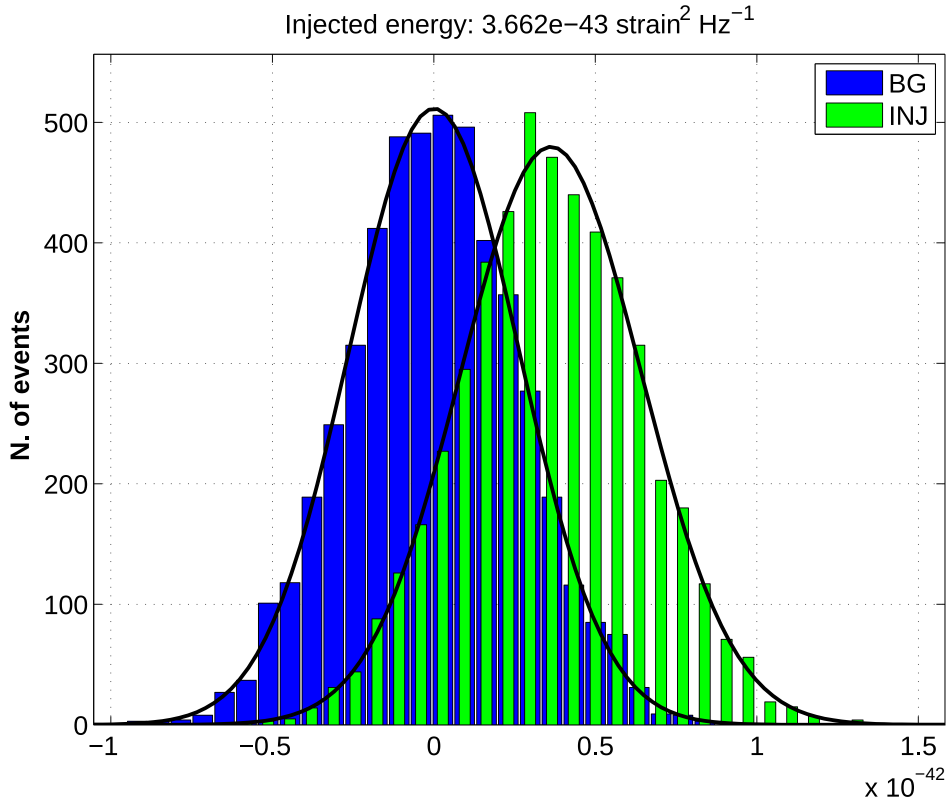

Figure 4. Simulated background and foreground distributions resulting from the search analysis pipeline using $M = 5000\ \Delta t$ long segments. The detector noise floor is described as white Gaussian with a strain equivalent noise of 10^{-22} strain Hz$^{-1/2}$, similar to the detector sensitivity at the time of the hyperflare event of 27 December 2004. For this simulation we assumed a QPO observation at 100 Hz and lasting for 50 s (similar to the observations in Ref. [11]). Due to this, the band center frequencies, bandwidths and signal durations are set to $f_{\rm qpo} = 100$ Hz, $f_- = 90$ Hz, $f_+ = 110$ Hz, $\Delta f = 10$ Hz and $\Delta t = 50$ s. The resulting (numerical) energy sensitivity, $E^{sens} = 3.66 \times 10^{-43} \pm 3 \times 10^{-45}$ strain^2 Hz^{-1} is found to be consistent within $\sim 1\%$ to the expected analytical sensitivity $E^{sens} = 3.63 \times 10^{-43}$ strain^2 Hz^{-1} of Eq.(19).

Carlo numerical results, as well as the background sensitivity studies using LIGO data (and presented in Fig. (3) of Ref. [20]), directly validates our search pipeline.

We simulated the detector noise as white Gaussian with a strain equivalent noise of 10^{-22} strain Hz$^{-1/2}$, similar to the H1 detector sensitivity at the time of the hyperflare event of 27 December 2004. Assuming a QPO observation at 100 Hz and lasting for 50 s (similar to the observations in Ref. [11]), we defined the time-frequency volume with band center frequencies, bandwidths and signal durations set to $f_{\rm qpo} = 100$ Hz, $f_- = 90$ Hz, $f_+ = 110$ Hz, $\Delta f = 10$ Hz and $\Delta t = 50$ s. The resulting numerical energy sensitivity, $E^{sens} = 3.66 \times 10^{-43} \pm 3 \times 10^{-45}$strain^2 Hz^{-1} is found to be consistent within $\sim 1\%$ to the expected analytical energy sensitivity $E^{sens} = 3.63 \times 10^{-43}$ strain^2 Hz^{-1} of Eq.(19). In terms of $h_{rss-det}$ (amplitude) sensitivity, the search algorithm presented here provides a sensitivity of

$$h_{rss-det}^{sens} = 6.06 \times 10^{-22} \text{ strain Hz}^{-1/2} \quad (22)$$

The algorithm presented here is based on an excess energy search, essentially an autocorrelation procedure. Multiple LIGO detectors were taking data at the recurrent flares originating from SGR $1806-20$ and SGR $1900+14$ during the fourth (S4) and fifth (S5) LIGO science runs. The algorithm can easily be adapted to measure a cross-correlation power in order to exploit multiple data streams.

Multiple LIGO detectors were taking data at the recurrent flares originating from SGR $1806-20$ and SGR $1900+14$ during the fourth (S4) and fifth (S5) LIGO science runs. The algorithm presented here is essentially an autocorrelation procedure and can easily be adapted to measure a cross-correlation energy in order to exploit multiple data streams.

6. Acknowledgments

We are indebted to many of our LIGO colleagues, in particular Jolien Creighton, Patrick Sutton and Zsuzsa Márka for helpful discussions. The authors are grateful for the support of the United States National Science Foundation under cooperative agreement PHY-04-57528 and Columbia University in the City of New York. The authors gratefully acknowledge the support of the United States National Science Foundation for the construction and operation of the LIGO Laboratory and the Particle Physics and Astronomy Research 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. The authors also gratefully acknowledge the support of the research by these agencies and by the Australian Research Council, the Natural Sciences and Engineering Research Council of Canada, the Council of Scientific and Industrial Research of India, the Department of Science and Technology of India, the Spanish Ministerio de Educacion y Ciencia, The National Aeronautics and Space Administration, the John Simon Guggenheim Foundation, the Alexander von Humboldt Foundation, the Leverhulme Trust, the David and Lucile Packard Foundation, the Research Corporation, the Alfred P. Sloan Foundation.

References

[1] Woods P M and Thompson C 2004 *ArXiv Astrophysics e-prints* (*Preprint* `astro-ph/0406133`)

[2] Mazets E P *et al.* 1979 *Nature* **282** 587–589

[3] Hurley K *et al.* 1999 *Nature* **397** 41–43 (*Preprint* `astro-ph/9811443`)

[4] Hurley K *et al.* 2005 *Nature* **434** 1098–1103 (*Preprint* `astro-ph/0502329`)

[5] Woods P M *et al.* 2007 *Astrophys. J.* **654** 470–486 (*Preprint* `astro-ph/0602402`)

[6] Duncan R C and Thompson C 1992 *Astrophys. J. Lett.* **392** L9–L13

[7] Schwartz S J *et al.* 2005 *Astrophys. J. Lett.* **627** L129–L132 (*Preprint* astro-ph/0504056)

[8] Palmer D M *et al.* 2005 *Nature* **434** 1107–1109 (*Preprint* astro-ph/0503030)

[9] Israel G L, Belloni T, Stella L, Rephaeli Y, Gruber D E, Casella P, Dall'Osso S, Rea N, Persic M and Rothschild R E 2005 *Astrophys. J. Lett.* **628** L53–L56 (*Preprint* astro-ph/0505255)

[10] Watts A L and Strohmayer T E 2006 *Astrophys. J. Lett.* **637** L117–L120 (*Preprint* astro-ph/0512630)

[11] Strohmayer T E and Watts A L 2006 *Astrophys. J.* **653** 593–601 (*Preprint* astro-ph/0608463)

[12] Duncan R C 1998 *Astrophys. J. Lett.* **498** L45 (*Preprint* astro-ph/9803060)

[13] Strohmayer T E and Watts A L 2005 *Astrophys. J. Lett.* **632** L111–L114 (*Preprint* astro-ph/0508206)

[14] Levin Y 2006 *Mon. Not. R. Astr. Soc.* **368** L35–L38 (*Preprint* astro-ph/0601020)

[15] Levin Y 2006 *ArXiv Astrophysics e-prints* (*Preprint* astro-ph/0612725)

[16] Abbott B *et al.* 2004 *Nuclear Instruments and Methods in Physics Research A* **517** 154–179

[17] Flanagan É É and Hughes S A 1998 *Phys. Rev. D* **57** 4535–4565 (*Preprint* gr-qc/9701039)

[18] Flanagan É É and Hughes S A 1998 *Phys. Rev. D* **57** 4566–4587 (*Preprint* gr-qc/9710129)

[19] Anderson W G, Brady P R, Creighton J D and Flanagan É É 2001 *Phys. Rev. D* **63** 042003–+ (*Preprint* gr-qc/0008066)

[20] The LIGO Scientific Collaboration *ArXiv Astrophysics e-prints* To be submitted to Phys. Rev. D (*Preprint* astro-ph/0703419)

[21] Abramowitz M and Stegun I A 1972 *Handbook of Mathematical Functions* (Handbook of Mathematical Functions, New York: Dover, 1972)

[22] Hamming R W 1983 *Digital Filters* (Prentice-Hall)