

## A fixed false alarm probability figure of merit for gravitational wave detectors

This content has been downloaded from IOPscience. Please scroll down to see the full text.

2014 Class. Quantum Grav. 31 085004

(<http://iopscience.iop.org/0264-9381/31/8/085004>)

View the [table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 194.94.224.254

This content was downloaded on 08/01/2015 at 12:23

Please note that [terms and conditions apply](#).

# A fixed false alarm probability figure of merit for gravitational wave detectors

M Was<sup>1,2,3</sup>, P Kalmus<sup>4</sup>, J R Leong<sup>1,2</sup>, T Adams<sup>5</sup>, N Leroy<sup>6</sup>,  
D M Macleod<sup>5</sup>, C Pankow<sup>7</sup> and F Robinet<sup>6</sup>

<sup>1</sup> Albert-Einstein-Institut, Max-Planck-Institut für Gravitationsphysik,  
D-30167 Hannover, Germany

<sup>2</sup> Institut für Gravitationsphysic, Leibniz Universität Hannover, D-30167 Hannover,  
Germany

<sup>3</sup> Laboratoire d'Annecy-le-Vieux de Physique des Particules (LAPP),  
Université de Savoie, CNRS/IN2P3, F-74941, Annecy-Le-Vieux, France

<sup>4</sup> Theoretical Astrophysics, California Institute of Technology, Pasadena, CA 91125,  
USA

<sup>5</sup> School of Physics and Astronomy, Cardiff University, Cardiff, CF24 3AA, UK

<sup>6</sup> LAL, Université Paris-Sud, IN2P3/CNRS, F-91898 Orsay, France

<sup>7</sup> Department of Physics, University of Wisconsin–Milwaukee, Milwaukee, WI 53201,  
USA

E-mail: [jonathan.leong@aei.mpg.de](mailto:jonathan.leong@aei.mpg.de)

Received 4 November 2013, revised 24 January 2014

Accepted for publication 25 February 2014

Published 21 March 2014

## Abstract

Performance of gravitational wave (GW) detectors can be characterized by several figures of merit (FOMs) which are used to guide the detector's commissioning and operations, and to gauge astrophysical sensitivity. One key FOM is the range in Mpc, averaged over orientation and sky location, at which a GW signal from binary neutron star inspiral and coalescence would have a signal-to-noise ratio (SNR) of 8 in a single detector. This fixed-SNR approach does not accurately reflect the effects of transient noise (*glitches*), which can severely limit the detectability of transient GW signals expected from a variety of astrophysical sources. We propose a FOM based instead on a fixed false-alarm probability (FAP). This is intended to give a more realistic estimate of the detectable GW transient range including the effect of glitches. Our approach applies equally to individual interferometers or a network of interferometers. We discuss the advantages of the fixed-FAP approach, present examples from a prototype implementation, and discuss the impact it has had on the recent commissioning of the GW detector GEO 600.

Keywords: figure of merit, astrophysical reach, gravitational wave, detector, observatory, burst analysis, non-stationarity

PACS numbers: 04.80.Nn, 07.05.Kf, 95.85.Sz

(Some figures may appear in colour only in the online journal)

## 1. Introduction

The operation of a worldwide network of the first generation of interferometric instruments designed to detect GWs has recently come to a close. This network included two LIGO [1] interferometers, Virgo [2], and GEO 600 [3]. No detection of GWs were made during its operation. Currently only GEO 600 is collecting science data, while the other detectors are undergoing a major upgrade to the second generation interferometer network.

Data from the first generation network has been used to search for GWs from a variety of astrophysical sources including isolated neutron stars, coalescence of neutron stars and/or black holes, and stellar collapse [4–6]. The expected detectable signal duration depends on the source. For non-axisymmetric rotating neutron stars GWs are continuous and nearly monochromatic. For the coalescence of binary compact objects—such as neutron stars or black holes—the chirp-like signals reside in the first generation detector bandwidth for up to tens of seconds. A supernova however may emit a burst of GW energy which is only a fraction of a second long. These signals might be detected in blind searches or in coincidence with astrophysical events detected by electromagnetic or neutrino observatories such as supernovae or gamma-ray bursts. Non-GW information about an astrophysical event enables more sensitive searches, especially when a search only uses a single GW instrument.

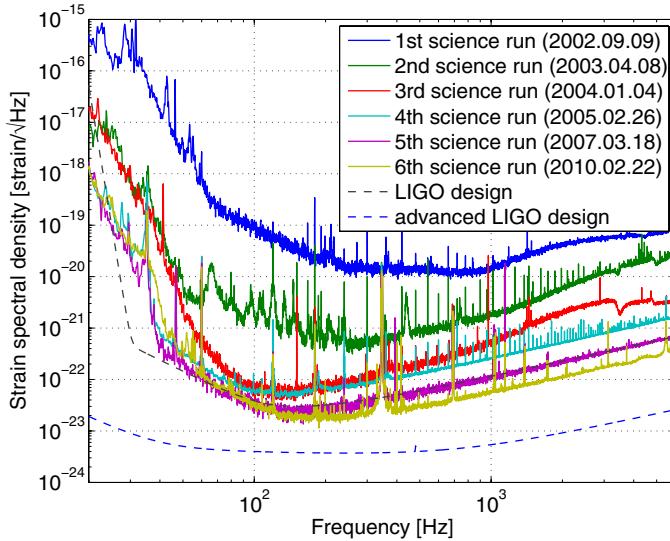
Characterizing the sensitivity of individual instruments, as well as networks of these instruments, to expected GW signals is important for a wide range of activities surrounding the GW discovery endeavor. These activities include the design of new instruments and networks, commissioning of the instruments, and reporting of scientific results.

On one hand, a simple astrophysical search sensitivity can be derived from estimations of the power spectrum of the interferometer output, assuming GWs do not contaminate the estimate appreciably. On the other hand, the ultimate measure of astrophysical sensitivity is carried out by the search itself, often run on a network of interferometers. In figure 1 we show example noise amplitude spectra from the various *science runs* (period of stable detector operations dedicated to astrophysical observations) of the 4 km LIGO interferometer located in Hanford, Washington. A simple sensitivity characterization, which was the primary figure of merit (FOM) for the detector operation during the running of the first generation network of interferometers, is the binary neutron star (BNS) range FOM [7, 8]. For first generation detectors the main commissioning goal was to improve and tune the instrument to increase that FOM.

The BNS range quantifies the canonical sensitivity of a search for GWs from the inspiral portion of the coalescence of two spinless  $1.4 M_{\odot}$  neutron stars on circular orbits. It estimates the range in Mpc at which such a GW signal would produce a matched-filter signal-to-noise-ratio (SNR) of 8 [7]. This range is volume averaged over the interferometer’s beam pattern and orbital inclinations of the binary relative to the line of sight. The range is given by

$$D_{\text{BNS}}(\rho) = \frac{K}{\rho} \left[ \int_{f_{\min}}^{f_{\text{ISCO}}} df \frac{1}{f^{7/3} S_n(f)} \right]^{1/2}, \quad (1)$$

where  $S_n(f)$  is an estimate of the one-sided power spectral density of the interferometer output,  $f_{\text{ISCO}} = c^3/(6\sqrt{6}\pi GM)$  is the frequency of GW emission at the innermost stable circular orbit



**Figure 1.** Progression of the 4 km LIGO Hanford detector noise spectra over the six science runs. These correspond to large increases in the BNS range in between the six science runs, respectively: 0.02 Mpc, 0.5 Mpc, 6.5 Mpc, 7.7 Mpc, 16 Mpc and 19 Mpc. Data obtained from the LIGO laboratory<sup>8</sup>.

in a Schwarzschild metric,  $K$  is a constant of proportionality involving only the binary mass and physical constants, and  $\rho$  is the chosen SNR threshold. In the following sections we describe how in general a range FOM is derived, and obtain the BNS range as a particular case. Here it is important to notice that the BNS range FOM is simply a reduction of the information contained in the noise power spectral density of an interferometer into a single number.

However the noise power spectral density does not completely characterize interferometer noise. Searches looking for short duration GWs found many transient events that were clearly instrumental artifacts such as photo-diode saturations or scattered light [9]. These *glitches* were found to limit the search reach by a significant factor over what the reach would be in the case of Gaussian instrument noise, especially for un-modeled GW burst searches. This limitation has been as large as a factor of two which results in a factor 8 smaller rate of observable events. In such cases the simple BNS range FOM described above does a poor job at estimating the search sensitivity because noise power spectral estimation is designed to characterize Gaussian noise. Essentially, it measures the noise standard deviation in different frequency bins, which depends mostly on the typical fluctuations of the noise and is relatively insensitive to rare large excursions, especially when the power of these transient excursions is spread over a large number of bins. Since short duration GW searches are looking for transient GW signals in the data, they are naturally affected by transient noise events and therefore can be used as a tool for characterizing non-Gaussian noise. Such characterizations are routinely done as part of the analysis procedure but are rarely presented as FOMs for others outside of the analysis community.

We describe a range based FOM which lies between a simple sensitivity estimate like the BNS FOM and a complete measure of network sensitivity given by the results of an actual search. It uses search procedures to include non-Gaussian components of the noise but is computed in a continuous manner to provide rapid feedback to the GW community.

<sup>8</sup> [http://www.ligo.caltech.edu/~jzweizig/distribution/LSC\\_Data/](http://www.ligo.caltech.edu/~jzweizig/distribution/LSC_Data/)

In section 2 we introduce a simplified formalism for discussing astrophysical search sensitivity of glitchy instruments and compare the old and new range FOMs within this formalism. In section 3 we discuss real FOMs and introduce an example prototype applied to GEO 600 data where we contrast the behavior of the two methods for characterizing noise. In section 4 we discuss how the range FOMs have informed the commissioning process and argue for the importance of the new range FOMs in the next years of GW research.

## 2. Formalism

In this section we introduce a formalism to characterize astrophysical search sensitivity of gravitational wave (GW) detectors. For clarity we neglect for the moment the shape or frequency make-up of the astrophysical signal, and concentrate on aspects related to making a detection of a transient GW signal with an instrument whose background noise contains many glitches.

### 2.1. Astrophysical range FOM

The question at the heart of the discussion presented in this paper is what it means for a signal to be minimally detected. This section gives the basic concept behind an astrophysical range. Let us define the amplitude of the GW emitted from a source at some nominal distance  $D_0$  to be  $h_{\text{emit}}$ . Then the amplitude  $h_{\text{det}}$  of the signal in an interferometer at a distance  $D$  from the source along the same line of sight is given by

$$h_{\text{det}} \propto \frac{D_0}{D} h_{\text{emit}}. \quad (2)$$

Now let us introduce a statistic that can pick out and order the strength of possible GW events in the detector noise [10, 11]. This statistic  $\rho$  has properties similar to an SNR such that

$$\langle \rho \rangle \propto \frac{h_{\text{det}}}{\sqrt{S_N}} \quad (3)$$

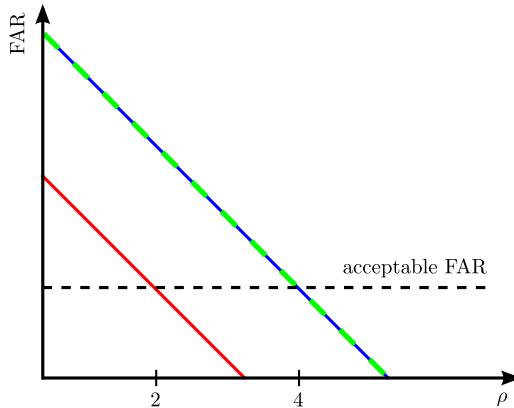
where  $S_N$  is the noise power spectral density at frequencies where the signal,  $h_{\text{det}}$ , also has spectral power. Here the proportionality factor depends on the specific statistic used. For the construction of the FOM we neglect the fluctuations of  $\rho$  due to noise and approximate it by its ensemble average  $\langle \rho \rangle$ . Combining relations (2) and (3) we can solve for the source distance corresponding to an event in the interferometer output characterized by the level  $\rho$  as

$$D = \frac{A h_{\text{emit}}}{\rho \sqrt{S_N}} \quad (4)$$

where  $A$  is the combined proportionality factors from relations (2) and (3). If we then take  $\rho$  to be the minimum value  $\rho_{\min}$  that we would claim as a detection of GWs from this type of source, then (4) represents the basic concept from which all astrophysical range FOMs are derived. This paper discusses different ways of choosing  $\rho_{\min}$  and what their consequences are in the cases of Gaussian and non-Gaussian noise.

### 2.2. Transient detector noise

Here we review some of the concepts involved in determining the significance of a candidate GW event in non-Gaussian noise [10–13]. Transient detector noise is characterized by creating a *background* data set in which the signals one is searching for are suppressed. When performing a transient GW analysis across multiple interferometers, the background data is typically created by shifting the time-streams of the individual interferometers by relative



**Figure 2.** Cartoon cumulative FAR histogram. At a given  $\rho$  this plots the FAR for all events with larger  $\rho$  values. The three histograms plotted here correspond to the three scenarios in figure 3. Here  $\rho = \frac{h_{\text{det}}}{\sqrt{S_N}}$ . We specify a maximum acceptable FAR for stating a detection shown as the dashed black line. The corresponding  $\rho$  thresholds,  $\rho_{\min}$ , which would be applied to foreground data for searches then specifies the amplitudes in figure 3 as  $h_{\text{det},\min} = \rho_{\min}\sqrt{S_N}$ .

offsets longer than both the light travel time between any two interferometers and the auto-correlation timescale of the searched signal. This ensures that there are no time coincidences in the data caused by a single transient GW event. Given that the expected rate of GWs is very low, chance coincidences between a time shifted GW signal and some other transient in the data does not bias noticeably the background data set. GW searches conducted in coincidence with other types of instruments, like gamma-ray detectors or optical telescopes, only search within a short temporal window around the observations of those instruments. Hence all other times are considered as background, this is especially relevant for single GW detector searches for which the background could not be estimated otherwise.

When a search algorithm is applied to this background data set it will find events that can then be ranked by a statistic like the  $\rho$  given in (3), whether the noise is Gaussian or not. Once this event ranking is carried out, the background can be characterized by collecting all events with  $\rho$  rankings greater than or equal to a given value and calculate a rate as a function of this  $\rho$  threshold. This is called a cumulative false-alarm rate (FAR) histogram. For any given  $\rho$  level it gives the rate of louder events that would be detected in the background data set which in principle only contains noise. We will see in a bit that the FAR statistic is the central piece of characterization information for any detection statement. Figure 2 shows an example of such a histogram.

After this characterization of the background data is done, a search is then carried out by running the same algorithm on a *foreground* data set. This is the data set that should contain GWs. In multiple detector analyses, data sets from the different detectors are matched in coincidence using GW travel times between sites. For external coincidence searches the foreground would be a short time surrounding the event time in the partner instrument. Normally, the background data sets contain much more data than the foreground, so that the background characterization has sufficiently low statistical errors.

Each event in the foreground is characterized by the  $\rho$  detection statistic. For any event the cumulative FAR histogram gives an estimate of the rate at which louder events due to instrumental glitches occur. Then, given the length of time spanned by the foreground data set and assuming a Poisson distribution of background events, this rate can be converted to a probability that the event is simply noise. This is the false-alarm probability (FAP) or p-value

characterizing the level of confidence that a given event is not due to a glitch. This value is then used to derive the confidence of a detection statement.

It is a matter of personal choice at what level of confidence one accepts the detection to be true. Nonetheless, in this paper we turn this process around by setting a given FAP as the threshold for claiming a detection. One can then use the expected data taking length and current knowledge of the cumulative FAR histogram to calculate a  $\rho_{\min}$  threshold above which one might claim a detection of GWs.

### 2.3. Choice of detection threshold

In section 2.1 we discussed how astrophysical range FOMs are based on (4) where  $\rho = \rho_{\min}$  is the detection threshold. We intend to display what this statement means in light of the relationship between detection significance and the FAP. Assuming the instrument noise is Gaussian, one can simply fix  $\rho_{\min}$  to a certain value,  $\rho_{\min} = \rho_{\text{fix}}$ . Using standard, Gaussian noise characterization techniques  $\rho_{\text{fix}}$  can be calculated from the desired FAP of a hypothetical search. The resulting *fixed-SNR range* would then be given by

$$D_\rho = \frac{Ah_{\text{emit}}}{\rho_{\text{fix}}\sqrt{S_N}} \propto \frac{1}{\sqrt{S_N}}. \quad (5)$$

We see here that  $D_\rho$  is simply inversely proportional to the square root of the noise power  $S_N$  and is thus a good measure of the noise floor, but not of other properties of the noise. For noise which contains glitches or other forms of short transient noise, the effect of these glitches on  $D_\rho$  is small compared to the effect it could have on the FAR because the measurement of  $S_N$  is meant to be robust against brief, transient fluctuations. In a phase of strong commissioning where the noise floor is changing by large amounts, as shown in figure 1,  $D_\rho$  functions to approximately tie the results of the work to astrophysical quantities. However, near the end of the commissioning process the Gaussian noise floor is evolving at a slower pace and the interferometers are preparing for long periods of science running operation where minimal commissioning is done. Here it is essential to have a FOM that more accurately mimics the analysis process. The downside as we will see in section 3.2 is that it comes with slow (order 1 h) reaction time scales.

At the end of section 2.2 we mentioned that a  $\rho_{\min}$  threshold could be measured from background data. Given a FAP threshold,  $\alpha_{\text{fix}}$ , and a length of foreground data,  $T$ , we can calculate a FAR threshold assuming Poisson statistics as

$$\gamma_{\text{fix}} = \frac{-\ln(1 - \alpha_{\text{fix}})}{T} \simeq \frac{\alpha_{\text{fix}}}{T}. \quad (6)$$

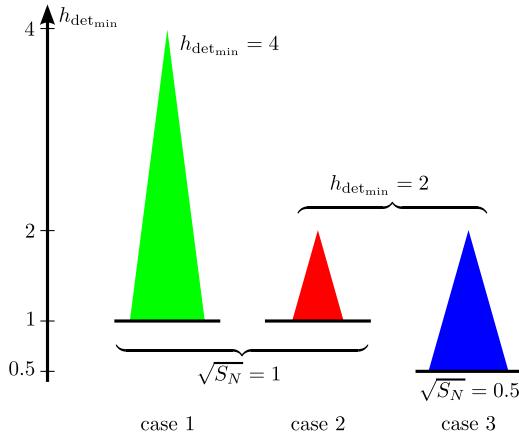
Representing the cumulative FAR histogram as  $\mathcal{F}(\rho)$ , we can then write the  $\rho$  threshold for detection as

$$\rho_{\min} = \mathcal{F}^{-1}\left(\frac{\alpha_{\text{fix}}}{T}\right). \quad (7)$$

This yields a *fixed-FAP range* based on (4) as

$$D_\alpha = \frac{Ah_{\text{emit}}}{\mathcal{F}^{-1}\left(\frac{\alpha_{\text{fix}}}{T}\right)\sqrt{S_N}} \propto \frac{1}{\mathcal{F}^{-1}\left(\frac{\alpha_{\text{fix}}}{T}\right)\sqrt{S_N}}. \quad (8)$$

We now see here that  $D_\alpha$  measures both the Gaussian noise floor through  $S_N$  and the transient noise fluctuations through  $\mathcal{F}^{-1}\left(\frac{\alpha_{\text{fix}}}{T}\right)$ . An alternative way to view (8) is to recall the



**Figure 3.** Pictorial representation of three different non-Gaussian noise scenarios. The black lines show the level of the Gaussian noise component, and the heights of the peaks the typical amplitude  $h_{\text{det},\min}$  of the non-Gaussian component. See figure 2 for the corresponding FAR histogram.

properties of  $\rho$  in (3) in combination with (7) and note

$$h_{\text{det},\min} \propto \mathcal{F}^{-1}\left(\frac{\alpha_{\text{fix}}}{T}\right) \sqrt{S_N} \quad (9)$$

$$\Rightarrow D_\alpha \propto \frac{1}{h_{\text{det},\min}}. \quad (10)$$

Here  $h_{\text{det},\min}$  is simply the minimum detectable  $h_{\text{det}}$  signal which takes into account the Gaussian noise floor as well as glitches.

Equation (8) is a simple toy version<sup>9</sup> of the main elements that go into the calculation of the exclusion distances or source rate-density upper limits which are the results of GW analyses [10–12, 14]. Because it mimics the analysis procedures, the fixed-FAP range accounts for glitches in the same way and thus gives a better representation of the output of an astrophysical search than a fixed-SNR range.

#### 2.4. Consequences of detection threshold choice

In order to illustrate the differences between the two choices of  $\rho_{\min}$  (fixed or tied to a FAP threshold), let us consider three cases where the instrument noise consists of some Gaussian noise as well as glitches. Figures 2 and 3 are a pictorial representations of the cases. The first two cases (from left to right) in figure 3 have the same level of Gaussian noise. The difference between the two is that the glitches in the second case are half as strong. In the third case, the glitches are just as strong as in the second case but the Gaussian noise floor is now half as high in amplitude. The corresponding FAR histograms are shown in figure 2 where we have chosen  $\rho = \frac{h_{\text{det}}}{\sqrt{S_N}}$ .

For these three cases we now calculate the fixed-SNR range,  $D_\rho$ , and the fixed-FAP range,  $D_\alpha$ , from (5) and (8) using  $A h_{\text{emit}} = 1$  and  $\rho_{\text{fix}} = 2$ . The results are shown in table 1. We see here that the fixed-SNR range is simply a measure of the Gaussian noise floor while the fixed-FAP range is a measure of the amplitude of the minimum detectable GW signal as was

<sup>9</sup> Compared to a real analysis we have neglected the fluctuations due to noise in the relation (3) between  $\rho$  and  $h_{\text{det}}$ . The effect of these fluctuation is measured in practice using a time consuming Monte Carlo.

**Table 1.** Value of fixed-SNR and fixed-FAP ranges for the three cases described in figures 2 and 3.

	Case 1	Case 2	Case 3
$D_\rho$	0.5	0.5	1
$D_\alpha$	0.25	0.5	0.5

alluded to in (5) and (10) respectively. For example, in going from case 1 to case 2 we see that as the minimum detectable signal decreases, our fixed-FAP range increases while the fixed-SNR range remains the same. This means that using a fixed-FAP range gives an incentive for reducing the transient background noise of the interferometers whereas the fixed-SNR range is blind to this effect. In going from case 2 to case 3 we see the opposite effect. The fixed-SNR range shows an increase as the noise floor is decreased but the glitch environment remains the same so the fixed-FAP range does not change. This time an incorrect incentive is given by the fixed-SNR range as a GW analysis would not have seen any difference between the two cases. Using a fixed-FAP range instead of a fixed-SNR range to monitor the commissioning process can tie successful commissioning more strongly to the possible science outcomes of the searches.

The next section discusses the issues that arise when attempting to create a realistic FOM using the simple toy model presented here as a basis. We also introduce and discuss an implementation of a fixed-FAP FOM.

### 3. Proof of concept

For most of the rest of this paper, we will restrict the discussion to short un-modeled GW sources or *bursts*. Due to the broad scope of these searches, they are more affected by instrumental glitches which are predominantly of short duration (sub-second). In comparison, GW searches looking for modeled signals that are more than a few seconds long are mostly unaffected by glitches once appropriate non-Gaussian noise rejection tests are used [15].

In this section, we describe an implementation of a fixed-FAP FOM as described in section 2 and give some examples of how such a FOM behaves with GEO 600 data in comparison to the standard fixed-SNR FOM. The FOM we have implemented here is only for a single instrument so for simplicity we also restrict the discussion mainly to single instrument FOMs. A discussion about the reasons for extending this to multiple instrument networks is presented in section 4. We limit our discussion to an FOM for burst searches, which in the past were most affected by glitches [9]. However, the formalism can be applied in a general manner.

#### 3.1. Realistic FOMs

We define a realistic FOM by first considering the frequency spectrum of the GW signal that would be seen in an interferometer. Let us postulate a signal of the form

$$h_{\text{det}} = \frac{k}{D} h_{\text{n}}(t) \quad (11)$$

where  $D$  is the distance between the source and the interferometer;  $k$  is a factor which depends on the emission amplitude, the direction to the source relative to the interferometer, and the

GW polarization; and  $h_n(t)$  is the normalized shape of the waveform. The normalization is such that

$$\int_0^\infty |H_n(f')|^2 df' = 10^{-42} \text{s} \quad (12)$$

where  $H_n(f)$  is the Fourier transform of  $h_n(t)$ , this is the order of magnitude of a waveform that can be currently detected by GEO 600.

When carrying out a matched filter analysis looking for the same waveform  $h_n(t)$  in the data output of the interferometer, whose noise spectral density is  $S_n(f)$ , the SNR for the signal (11) is given by

$$\rho = \frac{k}{D} \left[ \int_0^\infty \frac{|H_n(f')|^2}{S_N(f')} df' \right]^{\frac{1}{2}}. \quad (13)$$

As in section 2 we solve for  $D$  and substitute for  $\rho$  some minimum threshold SNR for declaring a detection. This gives a range FOM for a single source

$$D_{\min} = \frac{k}{\rho_{\min}} \left[ \int_0^\infty \frac{|H_n(f')|^2}{S_N(f')} df' \right]^{\frac{1}{2}}. \quad (14)$$

As discussed above, we can create either a fixed-SNR or fixed-FAP FOM simply by choosing  $\rho_{\min}$  differently. Note that the BNS range (1) is a particular case of (14) with the signal of the form  $|H_n(f)| \propto 1/f^{7/6}$  for  $f < f_{\text{isco}}$ .

As a simple concrete source model, we postulate a transient mass quadrupole rotating at a frequency  $\frac{f_0}{2}$  such that the resultant waveform is an enveloped sine waveform with central frequency  $f_0$  and duration  $\tau$ . An example of such a waveform is the sine-Gaussian:

$$h_{f_0\tau}(t + t_0) = h_0 \sin(2\pi f_0 t) e^{-\frac{t^2}{\tau^2}}, \quad (15)$$

where  $t_0$  is the time at the center of the waveform and  $h_0$  is an amplitude parameter. If the energy emitted from such a transient is given by  $E_{\text{GW}}$  then the average factor  $k$  assuming a distribution uniform in volume and source orientation [16] is

$$k(f_0) = A_0 \frac{1 \text{ kHz}}{f_0} \left( \frac{E_{\text{GW}}}{10^{-8} M_\odot c^2} \right)^{\frac{1}{2}}, \quad (16)$$

where  $A_0 \simeq 500 \text{ pc}$ . Now for  $\tau$  large enough such that  $S_N(f)$  does not vary significantly over the frequency range in which  $H_n(f)$  is non-negligible, we can simplify (14) by making use of the normalization condition (12) and substitute the amplitude  $k(f_0)$  to obtain

$$D_{\min}(f_0, \tau) = \frac{500 \text{ pc}}{\rho_{\min}(f_0, \tau)} \frac{1 \text{ kHz}}{f_0} \left( \frac{10^{-42} \text{s}}{S_N(f_0)} \frac{E_{\text{GW}}}{10^{-8} M_\odot c^2} \right)^{\frac{1}{2}}. \quad (17)$$

This is the minimum detectable range FOM for a single source waveform which is a short burst of energy  $E_{\text{GW}}$  and duration  $\tau$ , at frequency  $f_0$ . The unknown sky position and source rotation axis orientation has already been averaged over in the calculation of the amplitude factor  $k(f_0)$  in (16).

Un-modeled searches for GW bursts are meant to cover a wide range of short transients, which span the entire sensitive frequency band of the instruments and a range of durations. Thus an over-complete decomposition onto  $\mathcal{O}(1000)$  waveforms covering that parameter range is performed. Computational cost for the searches carried out during the era of the initial instruments was too prohibitive to allow the calculation of a minimum detectable range for each basis waveform. Instead, statistics on the backgrounds for a wide range of waveforms are collected together. When this is done, in place of an SNR detection threshold for each waveform  $\rho_{\min}(f_0, \tau)$  we set an overall threshold which is applied for all the waveforms

together. Let us call this threshold  $\rho_{\min}^{[f_1, f_2], [\tau_1, \tau_2]}$  where  $[f_1, f_2]$  and  $[\tau_1, \tau_2]$  simply label the frequency and duration range of the collected waveforms respectively.

In the end, since we do not know the details of the sources we are trying to search for it is useful to average over the possible frequencies creating one parameter, a marginalized range  $\langle D_{\min} \rangle$ , which gives information about the capabilities of the instrument for a given distribution of sources. In general, this distribution of sources could be a density

$$d(\mathbf{r}, f_0) \equiv \frac{dN_{\text{GW}}}{dt dV df_0}, \quad (18)$$

where  $\mathbf{r}$  is the position vector,  $N_{\text{GW}}$  is the number of transient events of energy  $E_{\text{GW}}$  and frequency  $f_0$ ,  $dt$  is the time element, and  $dV$  is the volume element. Let us consider the simplest case of a uniform distribution of sources in space and frequency so that  $d(\mathbf{r}, f_0) = \frac{B}{\Delta f_0}$  where  $B$  is simply the rate density of sources of all frequencies. Then the number of sources we expect to detect during a specified time interval of length  $T$  is

$$N_{\text{GW}} = \int d(\mathbf{r}, f_0) dt dV df_0 \quad (19)$$

$$= 4\pi T \frac{B}{f_2 - f_1} \int_{f_1}^{f_2} \int_0^{D(f_0)} r^2 dr df_0 \quad (20)$$

$$= \frac{4\pi}{3} \frac{TB}{f_2 - f_1} \int_{f_1}^{f_2} D(f_0)^3 df_0, \quad (21)$$

where  $D(f_0)$  is the single source range FOM (17) using the multiple source threshold,

$$D(f_0) = \frac{500 \text{ pc}}{\rho_{\min}^{[f_1, f_2], [\tau_1, \tau_2]}} \frac{1 \text{ kHz}}{f_0} \left( \frac{10^{-42} \text{s}}{S_N(f_0)} \frac{E_{\text{GW}}}{10^{-8} M_{\odot} c^2} \right)^{\frac{1}{2}}. \quad (22)$$

If we define

$$\langle D_{\min} \rangle \equiv \left[ \frac{1}{f_2 - f_1} \int_{f_1}^{f_2} D(f_0)^3 df_0 \right]^{\frac{1}{3}}, \quad (23)$$

we can easily calculate the number of expected events for a given uniform density source population as

$$N_{\text{GW}} = \frac{4\pi}{3} \langle D_{\min} \rangle^3 TB. \quad (24)$$

This is exactly the characteristic we would like to have from a marginalized range FOM for searches of un-modeled GW bursts coming from a uniform (in space and over a frequency interval  $[f_1, f_2]$ ) distribution of possible sources. Thus  $\langle D_{\min} \rangle$  defined in (23) and (22) is the form of the range FOM which we will deal with for the rest of this paper.

### 3.2. Determination of FAR threshold

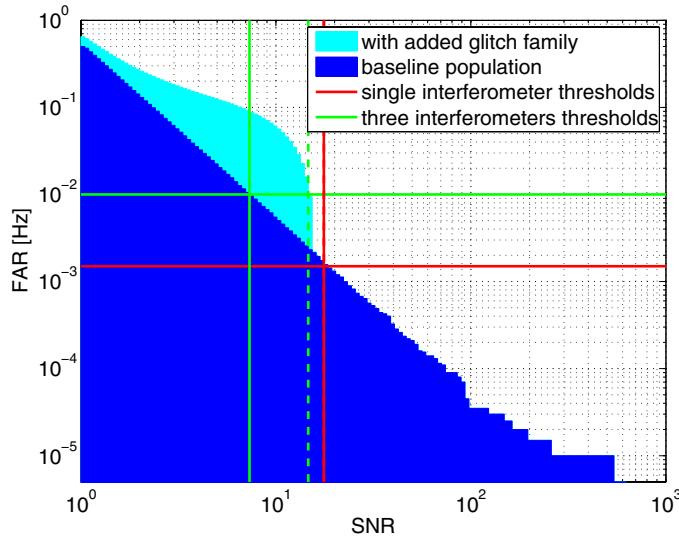
In the previous section we introduced the general form of a realistic, single instrument GW burst range FOM. What remains now is the choice of the threshold  $\rho_{\min}^{[f_1, f_2], [\tau_1, \tau_2]}$  used in (22). It follows from the discussion in section 2 that a fixed-SNR FOM can be trivially constructed by setting  $\rho_{\min}^{[f_1, f_2], [\tau_1, \tau_2]} = \rho_{\text{fix}}$ . For the fixed-FAP FOM the threshold can be given by something similar to (7). Here we discuss how the chosen relation between the FAP and  $\rho_{\min}$  depends on the type of analysis considered and how the resultant FOM behaves in the face of different distributions of glitches.

The FOM example which will receive the most attention in the rest of this article is designed to provide feedback on a search for GWs seen in a single interferometer triggered by a detection of neutrinos from a galactic core-collapse supernova. Because the time of flight of the neutrinos is expected to be the same as for the GWs, the search would only look for transients in the interferometer data within a window of about two seconds long. This already includes delays for most GW emission scenarios (see [17, 18] for recent reviews). For this simple example search we can easily construct a single interferometer fixed-FAP FOM by following the same procedure the search analysis uses. We choose a FAP level for claiming a detection,  $\alpha_{\text{fix}}$ , and combine this with the length of the foreground data,  $T = 2$  s, as specified in (6). Here we will use a 3 sigma detection threshold which gives a FAP of  $\alpha_{\text{fix}} = 3 \times 10^{-3}$ , resulting in a FAR threshold  $\gamma_{\text{SN}} \equiv \frac{\alpha_{\text{fix}}}{T} = 1.5$  mHz.

This trivial choice becomes somewhat arbitrary when ones attempts to create a fixed-FAP FOM which does not exactly follow the search analysis procedure but instead attempts to estimate its results using a simplified procedure. For example, one could construct a single interferometer FOM which estimates the capabilities of a three interferometer network. Here we imagine a hypothetical detector network made of three identical interferometers that have the same noise spectrum and FAR histogram. An analysis could be carried out on this network by simply searching for coincident events in the independent instruments. If the three interferometers are separated by distances of 10 ms light travel time, then coincidence would require that all three interferometers see the event within a single  $\delta t = 10$  ms time window. For Poisson statistics, the probability of a false-alarm in an interferometer with a FAR of  $\gamma$  within the time window  $\delta t$  is approximately  $\gamma \delta t$ . Thus the probability that all three interferometers see an event within a window of length  $\delta t$  is  $(\gamma \delta t)^3$ . Turning this back into a rate gives the FAR for coincident analysis on a network of three identical interferometers  $\gamma_{\text{net}} \simeq \gamma^3 \delta t^2$ . We can use this relation to define what FAR threshold is reasonable for fixed-FAP FOMs of single interferometers extended into a network of three similar interferometers. For a year long observation and a FAP of  $\alpha_{\text{fix}} = 3 \times 10^{-3} = \gamma^3 \delta t^2 \times 1$  year, we obtain  $\gamma = 10$  mHz  $\equiv \gamma_{3\text{det}}$ . This is the single interferometer FAR which, when combined with two other similar interferometers in coincidence for one year of operation, would produce a FAP of  $3 \times 10^{-3}$ .

We can see from these two examples with very different analyses and interferometer network configurations that a single FAP can result in very different threshold FARs for their respective FOMs. For these kinds of threshold based analyses, the corresponding FARs have a large effect on which populations of glitches affect the FOMs. To illustrate this effect let us consider an interferometer whose noise spectrum never changes so any fixed-FAP range FOM is simply inversely proportional to the SNR threshold determined from the cumulative FAR histogram and the FAR threshold. We begin with a base population of background events such that the cumulative FAR histogram is an inverse-squared function of the event SNR. This is plotted in figure 4 (blue histogram). The SNR thresholds,  $\mathcal{F}^{-1}(\gamma_{\text{fix}})$ , corresponding to the two FAR thresholds discussed are shown in table 2 for this and the following scenarios.

Now imagine that on top of this base population of glitches we add another population which consists of frequent, low amplitude glitches as shown in cyan in the figure. We see that this addition affects the three interferometer network analysis FOM with the high FAR threshold but has no effect on the single interferometer analysis FOM. This occurs because the glitches in the added population all have SNRs below the single interferometer analysis SNR threshold set by the base population. Since each bin in the cumulative histogram is constructed from glitches with SNRs greater than a given SNR, it does not matter how many glitches we add to the population below the threshold in some base population. For the three interferometer



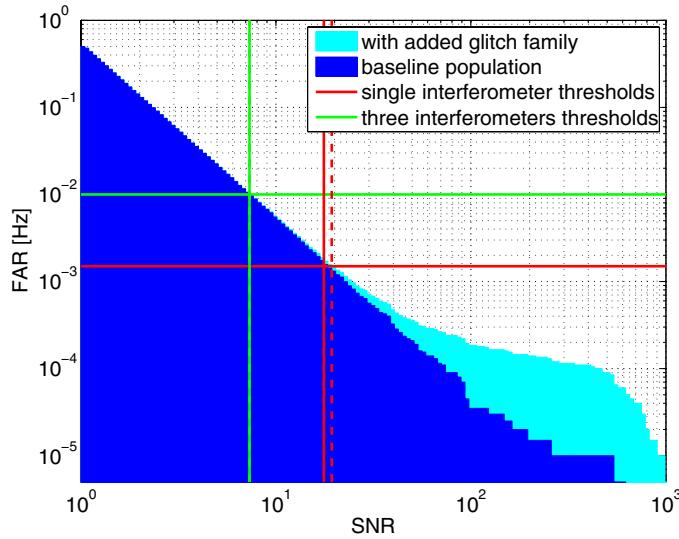
**Figure 4.** Cumulative FAR histogram of a distribution whose dependence on glitch SNR is inverse-squared (blue) where the total rate of glitches is 0.5 Hz. The horizontal red and green lines denote FAR threshold choices for FOMs tailored to two different analyses as described in this section (single and three interferometer network analysis respectively). The corresponding solid vertical lines show the resulting SNR thresholds,  $\mathcal{F}^{-1}(\gamma_{\text{fix}})$ , for this distribution. A resultant fixed-FAP range FOM would then be an inverse function of this SNR threshold as shown in (22) and (23). In cyan we show the cumulative FAR histogram for the same events with an added family of glitches having a high rate of 0.15 Hz and low SNR (uniformly distributed between 0 and 15). The SNR thresholds for this new distribution are shown as the dashed vertical lines.

**Table 2.** A comparison of threshold  $\rho_{\min}$  (and therefore inversely proportional to the fixed-FAP range FOM) for different glitch populations added onto a starting base population. The FAR thresholds used here correspond to FOMs tailored to two different analyses (1.5 mHz for a single interferometer search and 10 mHz for a three interferometer network search) as described in this section.

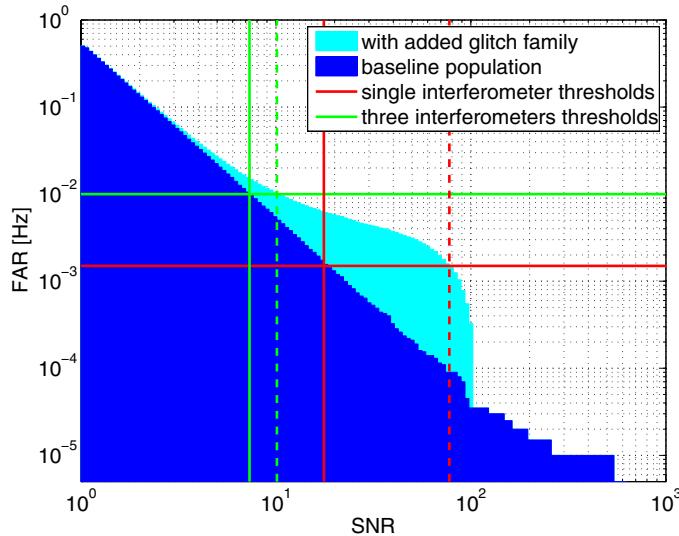
	$\mathcal{F}^{-1}(1.5 \text{ mHz})$	$\mathcal{F}^{-1}(10 \text{ mHz})$
base population	18	7.3
+ high rate, low SNR	18	15
+ low rate, high SNR	19	7.3
+ med rate, med SNR	78	11

analysis FOM the new threshold gets set near 15, to the largest SNR allowed in the added population.

On the other extreme, we can consider adding a population of very large amplitude glitches which are very infrequent (figure 5). Here we see that the three interferometer network analysis FOM is not affected at all while the single interferometer analysis FOM gets a only 6% increase in its SNR threshold. In this scenario we have added glitches whose SNRs are chosen from a uniform distribution from 100 to 1000. Some of these have SNRs approximately 60 times larger than the SNR threshold of the single interferometer analysis FOM and 140 times larger than the three interferometer analysis FOM. However, the rate of the added population is 10 times and 67 times lower than the single and three interferometer analyses' FAR thresholds respectively. Therefore, although the added glitches move the entire cumulative histogram up



**Figure 5.** Same as in figure 4 but with the added family of glitches having a low rate of  $150\ \mu\text{Hz}$  and high SNR (uniformly distributed between 100 and 1000).



**Figure 6.** Same as in figure 4 but with an added family of glitches having a medium rate of  $5\ \text{mHz}$  and medium SNR (uniformly distributed between 10 and 100).

in rates, the change is minuscule. Instead, the SNR thresholds are dominated by the baseline population of glitches which are happening at rates near the FAR thresholds. It is worth noting here that for these threshold analyses, it doesn't matter how loud these infrequent glitches are, only that they are above the previous SNR threshold. They could have even been much louder and the effect would have been exactly the same.

In the final scenario (figure 6), let us add an intermediate population of glitches to the base population with a rate lying between the last two additions. Since this population has glitches of larger amplitude than the SNR threshold and has a rate higher than the FAR threshold for

the single interferometer analysis FOM, we see that the SNR threshold behaves as the three interferometer analysis FOM did in the first scenario. It simply jumps to near the loudest allowable SNR in the additional population. The three interferometer analysis FOM behaves similarly to the single interferometer analysis FOM in the second scenario by showing a small increase. The proportional increases here are different from the previous scenarios because of the different relationships between the thresholds and the glitch family parameters, however the concepts are the same.

It should be clear from this investigation that a fixed-FAP FOM is only sensitive to the highest SNR family of glitches that happen at rates near or higher than the FAR threshold. Unless one displays multiple FOMs for different FAP levels, this method can potentially hide large amplitude backgrounds that could exist at lower FAP levels. Finally, it should be kept in mind that this was all done assuming the spectral noise estimation is not affected by the glitches. For a sufficiently large population of glitches the noise spectral estimation would rise and affect the range FOM as well.

### 3.3. Example implementation

We now describe an implementation of a fixed-FAP FOM. As mentioned in section 3.2, this is tailored to give feedback on searches for GW using single interferometer GEO 600 data triggered externally by the detection of neutrinos from a galactic core-collapse supernova. These supernovae are caused by the core-collapse of a massive star which is then followed by the formation and excitation of a proto neutron star with rotation frequencies predicted to be in the range of 250–1000 Hz. There are many different mechanisms for producing GWs from core-collapse supernovae [17, 18]. One could attempt to incorporate all of them into a range FOM, however for simplicity we will only consider here the rotational excitations of the proto neutron star. This would produce GWs within the frequency range 500–2000 Hz which is the most sensitive region of GEO 600. In the end, we will see that the calculated ranges will be smaller than the thickness of the disk of the Milky Way in the vicinity of the Solar System so we can assume that the possible sources are distributed uniformly in space.

Given this astrophysical scenario, we can then apply the treatment in section 3.1 to define a core-collapse supernova range as

$$D_{\text{SN}} \equiv \langle D_{\min} \rangle = \frac{500 \text{ pc}}{\rho_{\min}^{[f_1, f_2], [\tau_1, \tau_2]}} \left( \frac{E_{\text{GW}}}{10^{-8} M_{\odot} c^2} \right)^{\frac{1}{2}} \times \left[ \frac{1}{1.5 \text{ kHz}} \int_{0.5 \text{ kHz}}^{2 \text{ kHz}} \left( \frac{10^{-42} \text{ s}}{S_N(f)} \right)^{\frac{3}{2}} \left( \frac{1 \text{ kHz}}{f} \right)^3 df \right]^{\frac{1}{3}}. \quad (25)$$

For the power spectrum,  $S_N(f)$ , estimation we use the median averaging method [15] with one second long spectra averaged over one minute of data. These parameters are comparable to what has been used to compute the BNS range as displayed in the detectors' control rooms. A consequence of this averaging is that the spectrum is most affected by glitches that are about one second long which was the expected duration of BNS signals.

We will use equation (25) to define two range FOMs to compare. For a fixed-SNR FOM we will simply set

$$\rho_{\min}^{[f_1, f_2], [\tau_1, \tau_2]} = \rho_{\text{SN}} = 6.5. \quad (26)$$

This value is determined by calculating the threshold corresponding to a FAP of  $3 \times 10^{-3}$  assuming Gaussian noise.

For a fixed-FAP FOM we first need to collect statistics about transient events from a suitable data analysis algorithm. Here we will use the Omega pipeline [19, 20]. This search

algorithm has been used for generic GW burst searches [21, 22] and has also proven to be a useful tool for low latency detector characterization [9]. This makes Omega a convenient choice as the event generator for a fixed-FAP FOM.

Once we have the collection of events for a given time, we can create the cumulative FAR histogram  $\mathcal{F}(\rho)$  as described in section 2. For our fixed-FAP FOM, following sections 2.3 and 3.2, we use the threshold

$$\rho_{\min}^{[f_1, f_2], [\tau_1, \tau_2]} = \mathcal{F}^{-1}(\gamma_{\text{SN}}) = \mathcal{F}^{-1}(1.5 \text{ mHz}). \quad (27)$$

However, the FAR histogram  $\mathcal{F}(\rho)$  may change with time as the detector performance and environmental conditions change. This raises the question of which data duration to use for the estimation of  $\mathcal{F}(\rho)$ . In other words what time-averaging should be performed on the non-Gaussian component of the noise. Shorter windows are subject to greater statistical uncertainty on the estimate of the value of the rate while longer windows will be slow to react to changes happening in the interferometer. On one side, the shortest window length that can be used is set by the inverse of the FAR threshold since lower FARs cannot be probed by shorter times. For our supernova range this is  $\frac{1}{\gamma_{\text{SN}}} = 667$  s. On the other side, the upper limit for a reasonable window length is set by considering the analysis type and the treatment of background non-stationarity. For an externally triggered supernova search, between a few hours and a few days of background data would be used assuming that the data behavior is not changing significantly on this time scale. In the case of the coherent all-sky GW burst search, since the background needs to be estimated at a FARs of  $10^{-3} \text{ yr}^{-1}$  and lower, a few months of data is used from each interferometer for background estimation. The changes in data behavior are then partially taken into account by splitting the data set into different data quality categories [9].

The more important consideration is the role the FOM will ultimately play. For tracking commissioning changes, it is often beneficial to see immediately the effects of changes made to the instrumentation. However, if the FOM is intended for helping with the decision to start a science run, it then makes sense to use longer background estimation windows. In the study presented, a window length of 4000 s is used since the FOM was initially designed as a tool to aid the commissioning process. In the plots below, this window slides along the data every two minutes producing a new FOM data point which is highly correlated to the neighboring points.

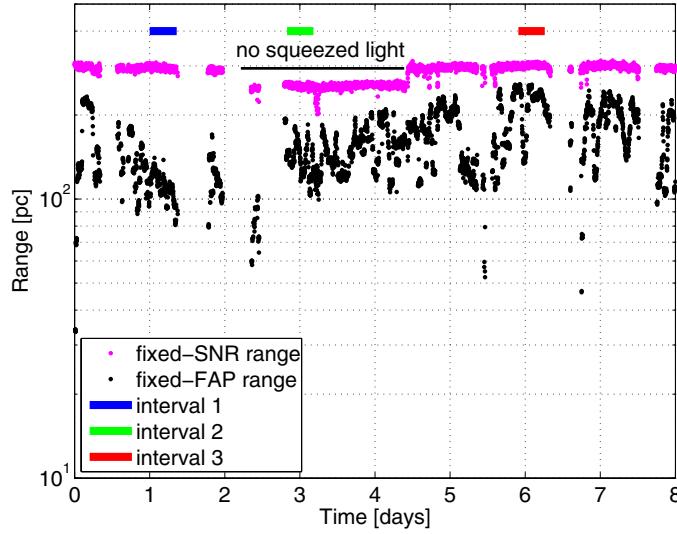
An easy way to construct the SNR threshold  $\mathcal{F}^{-1}(\gamma_{\text{fix}})$  can be seen by viewing it as the SNR for which the events of equal to or higher SNRs occur at a rate  $\gamma_{\text{fix}}$ . For window lengths that are integer multiples of the shortest possible length,  $n \times \frac{1}{\gamma_{\text{fix}}}$ , the SNR threshold simply becomes the SNR of the  $n$ th loudest event. Thus in the following section we will take as  $\mathcal{F}^{-1}(1.5 \text{ mHz})$  the sixth loudest glitch in 4000 s.

### 3.4. Behavior in real data

We now present the results of applying the FOMs described above on data from GEO 600. Figure 7 plots a comparison of the fixed-SNR and fixed-FAP supernova ranges as described in section 3.3 for eight days of data. There is only one configuration change of the interferometer where the squeezed light source [23] was switched off between days 2.3 and 4.4. The squeezed light source is used to decrease photon shot noise which is the dominant source of Gaussian noise above one kHz. This effect can be seen in figure 8.<sup>10</sup>

Note the significant difference in the behavior between the two FOMs in figure 7. While the fixed-SNR FOM remains stable the fixed-FAP range fluctuates more rapidly, changing by

<sup>10</sup> Another unknown effect happens to also be present here causing the noise floor to decrease in the several hundred Hz regime.



**Figure 7.** Comparison of fixed-SNR and fixed-FAP supernova ranges plotted over eight days of GEO 600 data. This data is not necessarily representative of long term GEO 600 sensitivity. The absolute scale is proportional to  $\sqrt{E_{\text{GW}}}$  and we assume here that supernovae emit  $E_{\text{GW}} = 10^{-8} M_{\odot} c^2$ . Three time intervals are chosen for closer inspection and are depicted by colored lines at the top of the figure: interval 1 is 8 h within (1.00, 1.36) days, interval 2 is 8 h within (2.83, 3.18) days, and interval 3 is 8 h within (5.91, 6.26) days. The black line denotes days 2.3 to 4.4 during which the squeezed light source was not operational.

a factor of several<sup>11</sup>. This time period is not necessarily representative of GEO 600 sensitivity but was chosen because it contains three time intervals with similar properties as the three scenarios presented in figure 3. Interval 1 is a case with high fixed-SNR range but low fixed-FAP range. Interval 2 is a case with lower fixed-SNR range but higher fixed-FAP range. The third interval has the same high fixed-SNR range as interval 1 and has the highest fixed-FAP range. These characteristics are summarized in table 3 and figure 10.

To get a clearer picture of what is happening here we take a look into the individual contributions to the FOMs. First we look at the spectral noise contribution by defining

$$\langle \sqrt{S_N(f)} \rangle \equiv \frac{\sqrt{10^{-42} \text{ s}}}{\rho_{\min}^{[f_1, f_2], [\tau_1, \tau_2]}} \frac{500 \text{ pc}}{D_{\text{SN}}} \frac{1 \text{ kHz}}{\langle f \rangle} \left( \frac{E_{\text{GW}}}{10^{-8} M_{\odot} c^2} \right)^{\frac{1}{2}} \quad (28)$$

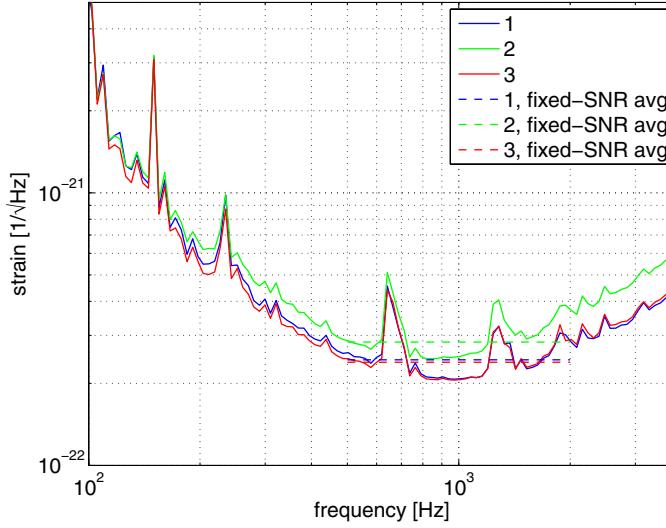
$$= \frac{1}{\langle f \rangle} \left[ \frac{1}{f_2 - f_1} \int_{f_1}^{f_2} \left( \frac{1}{S_N(f)} \right)^{\frac{3}{2}} \frac{df}{f^3} \right]^{-\frac{1}{3}}, \quad (29)$$

where

$$\langle f \rangle \equiv \left[ \frac{1}{f_2 - f_1} \int_{f_1}^{f_2} \frac{df}{f^3} \right]^{-\frac{1}{3}} \quad (30)$$

based on the range definition in (25). This is simply the particular type of frequency average of the spectral noise distribution which contributes to the two range FOMs. In figure 8 we

<sup>11</sup> This difference is larger than what one would expect between a fixed-FAP BNS range and the usual BNS range. The reason is that the Fourier transform duration used for power spectrum estimation here is much longer than the duration of expected GW burst signals while for BNS inspiral signals the durations are of the same magnitude.



**Figure 8.** Spectral noise estimation during time intervals defined in figure 7. The dotted lines show the spectral average (29) over the frequency range which is included in the SN range calculation.

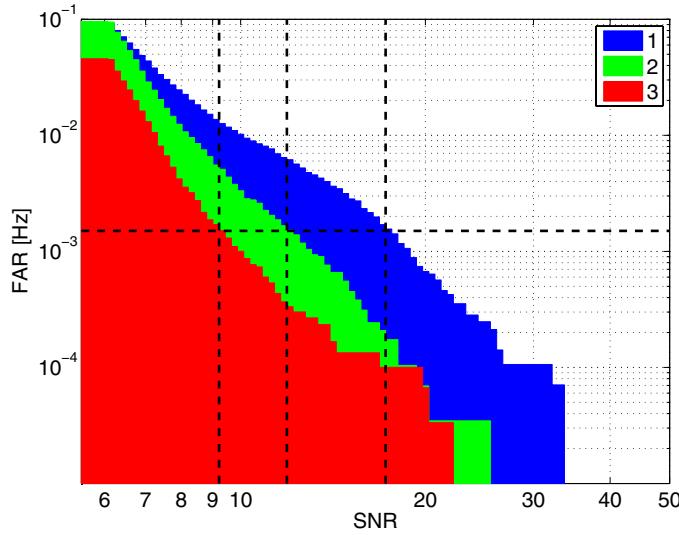
**Table 3.** GEO fixed-FAP SN range for the time intervals defined in the caption of figure 7. We also show here the two contributions to the range FOM which depend on the data: a measure of the Gaussian noise floor and a measure of the glitch distribution.

Interval	$D_{\text{SN}}$ [pc]	$\langle \sqrt{S_N(f)} \rangle [\text{Hz}^{-\frac{1}{2}}]$	$\mathcal{F}^{-1}(1.5 \text{ mHz})$
1	130	$2.4 \times 10^{-22}$	17
2	160	$2.8 \times 10^{-22}$	12
3	240	$2.4 \times 10^{-22}$	9.2

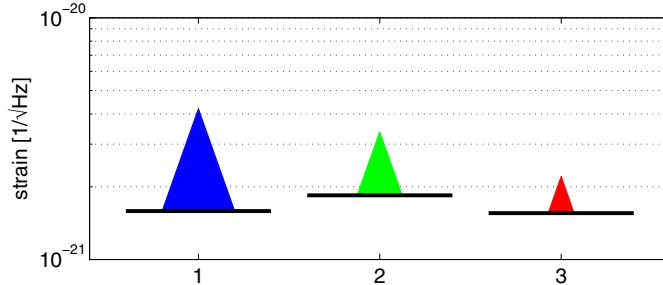
plot the average spectral noise level for the three intervals. Intervals 1 and 3 have very similar Gaussian noise characteristics while interval 2 has a spectral noise level which is 20% higher than the other intervals. As seen in figure 7, the squeezed light source was switched off during interval 2 which explains this behavior.

Let us now take a closer look at the non-Gaussian characteristics of the data. Figure 9 shows the cumulative FAR histograms for the 3 intervals. Here we use all the event triggers for each full 8 h interval. We see that the transient properties of the noise in each case are all very different. In particular, interval 1 and 3 have similar spectra but have cumulative FAR histograms with very different shapes. It is likely that the populations of glitches existing in the two intervals are very different from each other. In any case, this population change is not at all noticed by the spectral estimations but makes a large impact on the non-Gaussian measure,  $\mathcal{F}^{-1}(1.5 \text{ mHz})$ .

Finally in table 3 and figure 10, we combine the Gaussian and non-Gaussian properties of the data into the fixed-FAP range FOM and the corresponding  $\rho_{\min}$  threshold respectively. Here we see that in each of the cases studied here, the effect on the FAR-based range from the non-Gaussian nature of the data is far larger (85% between interval 1 and 3) and somewhat independent of the changes in the Gaussian noise floor (17% between interval 2 and the others). Thus the non-Gaussian properties dominate the movement of the range FOM. As mentioned in section 2, the fixed-FAP range incorporates the effect of non-Gaussian noise on a putative



**Figure 9.** Cumulative FAR histogram for intervals 1–3 defined in figure 7. Black dashed lines show the FAR corresponding to a FAP of  $3 \times 10^{-3}$  and the  $\rho_{\min}$  thresholds for the intervals shown on figure 7.



**Figure 10.** Pictorial representation of the minimum detectable signals,  $\rho_{\min}^{[f_1, f_2], [\tau_1, \tau_2]}(\sqrt{S_N(f)})$ , for the fixed-SNR (black lines) and fixed-FAP (tips of peaks) scenarios. The three cases here correspond to the three intervals defined in the caption of figure 7.

search and thus is more representative of an astrophysical search sensitivity than the fixed-SNR range.

#### 4. Discussion and conclusion

In the previous section we observed the markedly different behaviors of the fixed-SNR and fixed-FAP range FOMs in GEO 600 data. One way to describe the fixed-FAP range FOM is that it is a single number which integrates a combination of spectral sensitivity and glitch characterization in an interferometer or a network, in a way that mimics what is done in a search. As it is tailored for a specific search, a separate FOM for each type of astrophysical sources will be required. This new fixed-FAP range FOM should prove to be of increased benefit not only to predicting search performance, but also of particular use to diagnose

problems which arise during the construction and commissioning of various parts of the new second generation interferometers.

Due to their complexity, GW interferometers often require a large time investment in commissioning to make improvements. During a science run, operational time is very precious so commissioning must be optimized to reduce down time. The question then arises of how large of an effect any given improvement will have on the search capability? Also, what are the relative gains between having lower noise floor or less glitches? We have shown in this work that a fixed-FAP range in combination with a fixed-SNR version, as in figure 7, are good FOMs for answering questions like these. The fixed-SNR range gives information about the Gaussian noise floor, while the fixed-FAP range shows the effect of glitchy interferometer behavior. These FOMs are also useful throughout a science run as they can help with the monitoring of the interferometer status and its continuous fine tuning.

Currently on work days the GEO 600 interferometer is taken out of operation to commission upgrades or investigate noise sources. This operation is restored during nights and weekends. Since other interferometers are not operational, this provides for the opportunity to observe a serendipitously strong GW event if seen in coincidence with an exceptional astrophysical event. Due to this operational state, GEO 600 is being constantly tuned to produce the best possible science. This is why GEO 600 was a natural choice for prototyping a fixed-FAP range FOM. Some aspects of the interaction between the fixed-FAP supernova range and the commissioning process are worth discussion here.

At GEO 600 a study of the fixed-FAP range relative to the fixed-SNR range often caused the commissioning focus to shift strongly from investigating noise floor sources toward investigating glitch sources when the glitch background increased dramatically. Moreover, a fixed-FAP range forces glitch investigations to focus in a very natural manner on the glitches that are actually hurting the astrophysical range. For our supernova example these are the glitches which happen at a rate of  $5 \times 10^{-4}$  Hz. Other glitches which may have been more tempting to investigate without such a FOM, like the few loudest glitches happening every day or the most numerous, were labeled as second priority. After over a year of using the fixed-FAP supernova range in the GEO 600 control room, we have seen that it motivates the commissioning to proceed in a direction which has a better efficiency in improving the overall performance of the instrument relative to astrophysical searches. It is clear that such FOMs are essential for commissioning GW interferometers near or during a science run.

This work has focused on a single interferometer FOM meant to estimate the capability of a single interferometer GW search because GEO 600 is the only GW interferometer currently taking data. However, as construction completes on the Advanced LIGO, Advanced Virgo and KAGRA instruments, the GW detector network will again expand into a worldwide network [24]. It is thus important to further extend the concept described in this work to general networks of interferometers. In its most general form, this is to provide relatively quick and automatic feedback of the GW search capabilities of an interferometer or network of interferometers in the form of a FOM. This FOM could take on many different forms. On the simple end, single interferometer FOMs anticipating network searches could be created as described in section 3.2, however this should not impede the creation of full network FOMs which would most accurately estimate the search capabilities.

As an example, in the last network science run formed by the GEO 600 and Virgo interferometers during summer 2011, steps were taken to provide network analysis feedback to the commissioning work. Here parts of the coherent all-sky GW burst search procedure were run manually once a week on the network so that event backgrounds could be calculated. At this time GEO 600 had just finished the first implementation of a squeezed light source in a GW interferometer [23]. The squeezed light source decreased the Gaussian noise of GEO 600

by about 20% for frequencies above 1.5 kHz. Due to improper isolation of the injection path of the squeezed light there were a large amount (a factor of 10 more than nominal) of low amplitude glitches whenever the squeezed light source was operating. Since the science run had already started when this glitch excess was discovered, a simple decision to investigate the source of the glitches and eliminate them could not be made as this would use up a large portion of the science run time. We then turned to the network background events which were produced by the coherent all-sky GW burst search analysis to ask the question of how much this glitch excess is hurting the search. By investigating some data stretches with and without the squeezed light source in operation, we were able to see that the network SNR threshold that would be used by a search was increasing by no more than 10% when the squeezed light source was in operation. This feedback combined with the fact that the squeezed light source decreases the Gaussian noise floor by 20% gave us the confidence to continue running the squeezed light source throughout the science run without spending observational time to investigate the excess glitches.

This example illustrates the importance of getting feedback from GW search analyses into the interferometer control rooms. This kind of feedback is only possible with the experience that the GW research community has gained in analysis and computing during the era of the first generation network of GW interferometers. As the community walks into the era of the second generation network, it is important that we take steps toward strengthening the loop from the instruments to the analyses and back to the instruments. The stronger this loop is the more the system as a whole, a network of interferometers including the search analysis pipelines, can grow as a GW detection tool. The work presented here is one step in making the connection of the analysis pipelines to the instruments more apparent.

## Acknowledgments

We would like to thank the LIGO scientific and Virgo collaborations for providing a medium where stimulating discussions took place which ultimately were the motivating factors for this work. These collaborations have also allowed the use of GEO 600 GW data in the prototype FOMs displayed here. The Max Planck Society, Leibniz Universitt Hannover, the Science and Technology Facilities Council in the UK, the Bundesministerium fr Bildung und Forschung and the state of Lower Saxony in Germany, and the Volkswagen Foundation all made generous contributions which helped to make GEO 600 a reality. Many people over many years have also worked hard on the design and construction, as well as upgrades to GEO 600 that make it what it is today. We also acknowledge the institutes out of which the authors work, in particular the support provided to MW and JRL by the Max Plank Society. PK is a member of the LIGO Laboratory, supported by funding from the National Science Foundation. LIGO was constructed by the California Institute of Technology and the Massachusetts Institute of Technology with funding from the National Science Foundation and operates under cooperative agreement PHY-0757058. TA and DMM were supported by a studentship from the Science and Technology Facilities Council, in particular TA by the STFC Long Term Attachment ST/I505621/1. CP was supported by the NSF grant PHY0970074 and the UWM Research Growth Initiative.

## References

- [1] Abbott B P *et al* 2009 LIGO: the laser interferometer gravitational-wave observatory *Rep. Prog. Phys.* **72** 076901
- [2] Accadia T *et al* 2012 Virgo: a laser interferometer to detect gravitational waves *JINST* **7** P03012

- [3] Grote H *et al* 2010 The GEO 600 status *Class. Quantum Grav.* **27** 084003
- [4] Kokkotas K D 2008 Gravitational wave astronomy *Rev. Mod. Astron.* **20** 140
- [5] Sathyaprakash B S and Schutz B F 2009 Physics, astrophysics and cosmology with gravitational waves *Living Rev. Rel.* **12** 2
- [6] Andersson N *et al* 2013 The transient gravitational-wave sky arXiv:[1305.0816](#)
- [7] Finn L S and Chernoff D F 1993 Observing binary inspiral in gravitational radiation: one interferometer *Phys. Rev. D* **47** 2198
- [8] Sutton P J 2003 S3 performance of the ligo interferometers as measured by sensemonitor *Technical report LIGO-T030276*
- [9] Aasi J *et al* 2012 The characterization of Virgo data and its impact on gravitational-wave searches *Class. Quantum Grav.* **29** 155002
- [10] Abadie J *et al* 2012 All-sky search for gravitational-wave bursts in the second joint LIGO-Virgo run *Phys. Rev. D* **85** 122007
- [11] Abadie J *et al* 2012 Search for gravitational waves from low mass compact binary coalescence in LIGO's sixth science run and Virgo's science runs 2 and 3 *Phys. Rev. D* **85** 082002
- [12] Abbott B P *et al* 2010 Search for gravitational-wave bursts associated with gamma-ray bursts using data from LIGO Science Run 5 and Virgo Science Run 1 *Astrophys. J.* **715** 1438
- [13] Aasi J *et al* 2013 Search for gravitational waves from binary black hole inspiral, merger, and ringdown in ligo-virgo data from 2009–2010 *Phys. Rev. D* **87** 022002
- [14] Biswas R, Brady P R, Creighton J D E and Fairhurst S 2009 The loudest event statistic: general formulation, properties and applications *Class. Quantum Grav.* **26** 175009
- [15] Babak S *et al* 2013 Searching for gravitational waves from binary coalescence *Phys. Rev. D* **87** 024033
- [16] Sutton P J 2013 A rule of thumb for the detectability of gravitational-wave bursts arXiv:[1304.0210](#)
- [17] Ott C D 2009 The gravitational-wave signature of core-collapse supernovae *Class. Quantum Grav.* **26** 063001
- [18] Kotake K 2013 Multiple physical elements to determine the gravitational-wave signatures of core-collapse supernovae *C. R. Physique* **14** 318
- [19] Chatterji S 2005 The search for gravitational wave bursts in data from the second LIGO science run *PhD Thesis* Massachusetts Institute of Technology
- [20] Rollins J G 2011 Multimessenger astronomy with low-latency searches for transient gravitational waves *PhD Thesis* Columbia University
- [21] Abadie J *et al* 2010 All-sky search for gravitational-wave bursts in the first joint LIGO-GEO-Virgo run *Phys. Rev. D* **81** 102001
- [22] Abbott B P *et al* 2009 Search for gravitational-wave bursts in the first year of the fifth LIGO science run *Phys. Rev. D* **80** 102001
- [23] Grote H *et al* 2013 First long-term application of squeezed states of light in a gravitational-wave observatory *Phys. Rev. Lett.* **110** 181101
- [24] Aasi J *et al* 2013 Prospects for localization of gravitational wave transients by the advanced LIGO and advanced Virgo observatories arXiv:[1304.0670](#)