Contents

Introduzione

La teoria della Relatività Generale predice l'esistenza delle Onde Gravitazionali (OG), perturbazioni dello spazio-tempo generate da movimenti di masse almeno di ordine quadrupolare. La relatività prevede che l'energia delle OG generate da masse in movimento sia una piccolissima frazione della massa della sorgente e che la loro interazione con la materia sia estremamente debole. Questi aspetti sono contemporaneamente vantaggiosi e svantaggiosi. Il maggior svantaggio consiste nell'impossibilità di produrre OG in laboratorio, in quanto necessiteremmo di masse ed energie tipiche di sorgenti astrofisiche.

Tuttavia, la debole interazione delle OG con la materia fa sì che le stesse OG mantengano inalterate le informazioni sulla sorgente generatrice, contrariamente a quanto accade per la radiazione elettromagnetica. Sfortunatamente, il rate atteso di eventi nell'universo strettamente vicino alla terra è così basso che una possibile rivelazione è alquanto difficile.

La comunità scientifica ha sviluppato strumenti sofisticati allo scopo di rivelare le OG generate nel vicino universo. Tuttavia, la rivelazione è resa alquanto difficile dalla presenza di una vasta varietà di sorgenti di rumore: dagli inevitabili rumori strumentali (termico, elettronico, ecc...) a quelli ambientali (temporali, movimenti sismici, passaggio di aeroplani, attività umana nelle vicinanze, etc...).

Tutti questi fattori contribuiscono alla complessità di rivelazione di un'OG. Esiste, tuttavia, una prova indiretta della loro esistenza, fornita nel 1993 da R. Hulse e J. Taylor, i quali analizzarono il periodo orbitale del sistema binario composto da due pulsar PSR1913+16. La diminuizione del periodo osservata è, in effetti, in ottimo accordo con la generazione di OG prevista dalla Relatività.

Il primo tentativo di rilevare un'OG proveniente dallo spazio è stato realizzato da J. Weber nel 1957. Il suo prototipo di rivelatore si basava sulla misura delle oscillazioni di una barra cilindrica di alluminio, convertite poi in segnali elettrici mediante un trasduttore. Era il primo rivelatore acustico a barra risonante. In seguito, vennero progettati gli interferometri: rivelatori che utilizzano il principio dell'interferenza della luce per rivelare il passaggio dell'onda. Questo tipo di rivelatori presenta caratteristiche migliori di quelli acustici: una sensibilità maggiore ad una banda più larga di frequenza. Al giorno d'oggi, esistono cinque interferometri

operativi sul pianeta.

Virgo è situato a Pisa, in Italia, ed è un progetto comune di CNRS e INFN, gestito dalla Collaborazione Virgo dell'European Gravitational Observatory [1].

Il Laser Interferometer Gravitational-Wave Observatory (LIGO) è una collaboratione Caltech-MIT supportata dalla National Science Foundation [2]. Esso si compone di tre rivelatori dislocati negli USA: il primo (L1) presso Livingston, Lousiana, e gli altri due (H1 e H2) presso Hanford, Washington; e di un rivelatore (GEO) situato ad Hannover, in Germania, il quale è gestito da una collaborazione Inglese e Tedesca [3].

Ultimo rivelatore funzionante è TAMA, situato vicino a Tokyo, in Giappone [4]. Tuttavia la sensibiltà degli ultimi due non è comparabile con gli altri, per cui gli ultimi risultati scientifici utilizzano solamente i dati dei LIGO americani e di Virgo, secondo l'accordo stipulato dalle due collaborazioni iniziato con il quinto run scientifico di LIGO (S5) e il primo di Virgo (VSR1). S5 ha occupato il periodo da Novembre 2005 fino a Settembre 2007, con i rivelatori continuamente in presa dati e alla sensibilità progettata. Virgo si è unito in Maggio 2007, con una sensibilità comparabile a quella degli altri tre nella banda di alta frequenza (Sec. 2).

Il coinvolgimento di più rivelatori aumenta l'affidabilità di rivelazione di OG rispetto all'utilizzo di un singolo strumento. Per questo motico, i gruppi di Analisi Dati hanno sviluppato algoritmi di ricerca che si possono applicare sui dati di più rivelatori. Un esempio è **Waveburst**, che illustreremo in dettaglio più avanti (Sec 3). In questa tesi ci concentremo sull'Analisi dei Dati volta a ricercare OG di breve durata (inferiore al secondo), chiamati **Bursts**, in particolare sulla banda di High Frequency, ossia sopra i 1280 Hz.

Nel primo capitolo introdurremo la fisica delle OG, partendo dalla trattazione della Relatività Generale fino ad elencare una lista di sorgenti astrofisiche di maggiore interesse per il lavoro trattato. Continueremo nel secondo capitolo descrivendo i rivelatori di tipo interferometrico. Nel terzo capitolo tratteremo una breve introduzione della ricerca di tipo Bursts, per passare poi ad una descrizione dettagliata dell'algoritmo Waveburst. Esso si basa sul linguaggio ad oggetti C++, e permette di eseguire analisi complesse sui dati dei rivelatori con un utilizzo limitato di risorse macchina, specialmente riguardo alla memoria utilizzata e al tempo di calcolo. Nelle appendici vengono riportate alcune caratteristiche tecniche. Lo studio dell'algoritmo in tutte le sue parti ha richiesto un grande sforzo, che tuttavia ha permesso di comprenderne le capacità e prestazioni, con lo scopo di procedere ad una revisione del codice e di implementare cambiamenti o nuovi metodi con lo scopo di migliorarne le prestazioni stesse.

Il quarto capitolo tratta delle modifiche implementate al codice per migliorare la

ricostruzione della direzione dell'onda incidente, ovvero della sorgente generatrice. Questo è un lavoro ancora in corso d'opera, in quanto l'obiettivo prefissato non è ancora stato raggiunto. Per questo progetto sono state necessarie modifiche sostanziali alla parte di codice riguardante la ricostruzione della posizione, nonché la creazione di strumenti o figure di merito atte allo scopo di caratterizzare le prestazioni dell'algoritmo su questo aspetto.

L'ultimo capitolo riporta i risultati dell'applicazione dell'algoritmo sui dati reali della rete di rivelatori alla ricerca di possibili rivelazioni. La procedura utilizzata, chiamata *blind search*, consiste nell'implementazione di un'analisi statistica rivolta ad ottenere una stima sulla presenza di un possibile candidato all'interno dei dati. Questa analisi statistica comprende uno studio dettagliato sui dati dei rivelatori coinvolti per caratterizzare le prestazioni di rivelazione dell'algoritmo e il livello di rumore dei rivelatori durante il periodo analizzato. Questo studio preliminare permette di stabilire delle regole di decisione sull'eventuale presenza di un candidato, il cui risultato apre due possibili scenari: in caso di presenza, uno studio più accurato applicato ai dati per determinare la confidenza di rivelazione; oppure, in caso di assenza, il calcolo degli upper limits sulle forme d'onda testate.

L'ultimo capitolo contiene il riassunto dei risultati ottenuti e una breve trattazione sulle questioni aperte sui progetti in corso, concludendo con un accenno ai progetti futuri che nascono dagli sviluppi della collaborazione negli ultimi mesi.

Introduction

The theory of General Relativity (GR) predicts the existence of Gravitational Waves (GW), perturbations of the space-time metric generated by mass motions of quadrupolar order. The theory predicts that the energy of a GW generated from a certain process is a very small fraction of the source mass and that their interaction with the matter is extremely weak. These aspects give both advantanges and disadvantages. The most obvious disadvantage is that we cannot produce GW in laboratories because we would need the typical amount of mass and energy characterizing astrophysical objects.

However, as a conseguence of such weak interaction with matter, GWs contain inalterate informations on the source, contrarily to what happens with electromagnetic radiation. Unfortunately, the rate of GW generating processes in the nearby universe is so low that detections are difficult.

The scientific community has developed sophisticate instruments to detect GW from the nearby universe. However, these instruments are affected by a lot of noise sources: infact, to the unavoidable instrumental noises (such as electrical, thermal, etc..) due to the complicated design, we have to add environmental noises, generated by external sources, like bad weather, airplanes, seismic and human activity, etc...

This explains why GW detection is so difficult. An indirect proof of their existence has been given by R. Hulse and J. Taylor, in 1993, analysing the orbital period of the binary pulsar PSR1913+16. This observed decrease is consistent with the predicted generation of gravitational waves.

The first attempt to detect GW from the universe was performed by J. Weber, in 1957. The detecting principle was based on the measure of the oscillations of an aluminum cyilinder, which were then converted in electrical signals by a transductor. This was the first proyotype of the acoustic detectors. During following years, the community developed alternative detectors based on the light interference: such devices are called interferometers, and they are characterized by a larger band and a better sensitivity with respect to the acoustic ones. At the present time, there are five operating interferometers on the Earth.

The Virgo detector is a joint project of the CNRS and INFN, operated by the Virgo Collaboration at the European Gravitational Observatory [1], located near Pisa (Italy).

The Laser Interferometric Gravitational-Wave Observatory (LIGO) is a joint Caltech-MIT project supported by the National Science Foundation [2]. It is composed by four interferometers; three situated in the USA: one (L1) in Livingston, Lousiana and two (H1 and H2) in Hanford, Washington. The last one (GEO) is located in Hannover, Germany and it is a British-German collaboration [3].

The last detector is the Japanese TAMA, located near Tokyo, Japan [4]. Since, the current sensitivity of the last two mentioned detector is not comparable with the others, the latest scientific observational results are based on the jointly analysis involving the LIGO and Virgo interferometers. This agreement started with the fifth LIGO scientific run (**S5**) at the first Virgo run (**VSR1**). S5 run took place between November 2005 and September 2007 in continuos data-taking mode and designed sensitivity. The VSR1 run joined in May-September 2007, with a sensitivity comparable to LIGO detectors in the high frequency band (Sec. 2).

The use of multiple detectors increase the reliability of GW detection respect to a single one. Data Analysis groups have developed search algorithms based on the joint analysis of data from multiple detectors. One example of these algorithms is **Waveburst**, which will be illustrated in detail in this thesis (Sec. 3). Data Analysis is performed by four LIGO-Virgo search physics groups which focus on different targets: Bursts, Compact Binary Coalescences, Continuous Waves and Stochastic Background. In this thesis we concentrate on the first one, the **Burst** search, in particular on the so-called **High Frequency** band, i.e. over 1280 Hz.

The first chapter is a brief introduction on the GW physics, from the GR theory to a list of interesting astrophysical sources. The second chapter describes the main characteristics of GW detectors, in particular the interferometers. After a brief introduction on the Burst Search characteristics, the third chapter describes in details the framework of Waveburst. Waveburst is an object oriented code based on C++ language. Its structure allows to perform complicate analyses on detector data streams, reducing the computational load in time and memory. Some technical characteristics are reported in the Appendices. A great effort has been devoted to understand all its capabilities and performances. The goal of this work is to revise the code (from bugs, etc...) and implement new features and methods to improve the performances of the algorithm.

The fourth chapter describes the implemented modifications of the code to improve the calculation of the sky direction of the source generating the GW. Such work has not yet been completed, because the established goals were not reached. A big effort was dedicated to the code part involved with the coordinate

reconstruction. For this work it has been also necessary the implementation of instruments and adapted figures of merit to characterise algorithm performances.

The last chapter regards the application of the algorithm to real data from a network of interferometric detector, searching for possible detections. We adopted a *blind search* procedure, i.e. an implementation of a statistical analysis fitted to make a confidential decision on the presence of a possible candidate in the analyzed data. This statystical analysis includes a deep study on the detectors data to characterize the algorithm detection performances and the data noise level. This preliminary study allows to establish the decision rules to declare a possibile presence of an interesting candidate. The possible presence decide the following steps: a more accurate study of the data to characterise the confidence of the found candidate, or the calculation of the upper limits on the tested GWs.

Finally, in the conclusions we resume the results we have reached up so far and the open issues on the on-going works and on the future projects that we have been developing in the last months.

Chapter 1

A brief introduction on Gravitational Waves

Gravitational Waves (GW) are perturbations of the space-time metric predicted by the theory of General Relativity. We examine in this paragraph the analythical derivation of GW from Einstein's equation and we list the astrophysical sources of interest for the existing detector on the earth.

1.1 General Relativity

Starting from the assumptions that any arbitrary coordinate system can be used to describe the equations of physics and that, passing from one frame to another the coordinate definitions and the metric tensor change but the space-time distance remains invariate, it follows that the light velocity in vacuum is constant for all coordinate frames (also assumed in the Special Relativity) and that there is a relation between matter and space-time curvature. This relation is matematically described by the Einstein's equations:

$$R_{ij} - \frac{1}{2} R g_{ij} = \frac{8\pi G}{c^4} T_{ij}, \tag{1.1}$$

where R_{ij} is the Ricci tensor, g_{ij} is the metric tensor, R the Rienmann scalar and T_{ij} the energy-momentum tensor. The left part of this equation refers to the space-time curvature (and it is usually called Eintein tensor G_{ij}), while the right part is related to the energy distribution.
These equations are invariant, i.e. same form in each coordinate frame, and are second order equations, being the Ricci tensor and the Rienmann scalar linear combinations of first and second derivative of the metric tensor.[1] They are there-

[1] The Riemann scalar is the contraction of the Rienmann tensor: $R^i_{jkl} = \frac{\partial \Gamma^i_{jk}}{\partial x^m} - \frac{\partial \Gamma^i_{jm}}{\partial x^k} + \Gamma^i_{ks}\Gamma^s_{jl} -$

fore difficult to solve, it is better to introduce approximations: in this case, we use the *linear approximation*:

$$g_{ij} = \eta_{ij} + h_{ij} \tag{1.2}$$

where η_{ij} is the Minkowski's metric and h_{ij} is a small perturbation ($|h_{ij}| << 1$). To simplify expression, we define the following tensor:

$$\tilde{h}_{ij} = h_{ij} - \frac{1}{2}\eta_{ij}h \tag{1.3}$$

($h = g^{ij}h_{ij}$), and we use a convenient gauge.

A gauge choice is equivalent to a change of coordinate frame. A coordinate change is given by the relationhship: $x'_i = x_i + \xi_i$, where x'_i and x_i are respectively the new and old coordinates. Under this coordinate change, a generic tensor T is trasformed in the way:[2] $T'_{i,j} = T_{i,j} - 2\partial_{(i}\xi_{j)}$.
In this case it is convenient to use the Lorentz Gauge, defined by the following uguagliance:[3]

$$\tilde{h}'^{l}_{k,l} = 0 \tag{1.4}$$

So the Einstein tensor can be written:

$$G_{ij} = -\frac{1}{2}\Box\tilde{h}'_{ij} \tag{1.5}$$

Additionally,e assumption $T_{i,j} = 0$, i.e. in the vacuum, the equations become:

$$\Box\tilde{h}'_{ij} = 0 \tag{1.6}$$

where we consider the case far from the source, i.e. in the vacuum, so the energy momentum has null component.

We can write solutions of the last equation as a linear waves combination:

$$\tilde{h}'_{ij} = Re\int d^3k A_{ij}(\mathbf{k})e^{i(\mathbf{k}\cdot\mathbf{x}-\omega t)} \tag{1.7}$$

where we have introduced two quantities:

- k_i, the four dimensions vector that describes the wave: $k_i = (\omega, \mathbf{k})$
- A_{ij}, a tensor that does not depend on the space-time, but only on the space vector $\mathbf{k}$

$\Gamma^i_{ls}\Gamma^s_{jk}$. Γ^i_{jk} are the Christoffel's symbols: $\Gamma^i_{jk} = \frac{1}{2}g^{is}\left(\frac{\partial g_{js}}{\partial x^k} + \frac{\partial g_{ks}}{\partial x^j} - \frac{g_{jk}}{\partial x^s}\right)$

[2]The round brackets indicate symmetric sum: $T_{(ij)} = T_{ij} + T_{ji}$

[3]The comma simbol is used to identify the partial derivative: $(\cdot)_{,l} = \partial_l\cdot$

The Lorentz Gauge assures that these quantities satisfy the equation $k^i A_{ij} = 0$, which is the transverse property of the wave.
From the previous considerations, we realize that the tensor $\tilde{h}'_{ij}$ has only two indipendent components,and we can choose the coordinate frame such that:

- $\tilde{h}'_{0j} = 0$: null time component
- $\tilde{h}'^{i}_{i} = 0$: null trace

This is called the **Transverse-Traceless Gauge** (TT).
Usually the two indipendent components are described in the following way. If we consider a frame where the wave propagates on the z-axis, $\tilde{h}'_{ij} = \tilde{h}'_{ij}(t-z)$, and from transverse propriety $\partial^z \tilde{h}'_{zi} = 0$ we have that $\tilde{h}'_{ij}(t-z)$ is constant and we can assume that this constant is null because, when the distance tends to infinite, the components of the tensor vanish.
We therefore have that the only non-trivial components are: h_{xx}, h_{xy}, h_{yx}, h_{yy}. From Traceless assumption $h_{xx} = -h_{yy} \equiv h_+$ and from simmetry of the tensor: $h_{xy} = h_{yx} \equiv h_\times$.
These two quantities (h_+ and $h_\times$) are the two **wave polarizations**.

1.2 Astrophysical Sources

We have introduced the linear approximation to the Einstein equation and used the Lorentz Gauge to write in a simpler way the Einstein tensor (Eq. 1.5). In the previous paragraph we have considered the case of a point far from the source, i.e. the energy-momentum tensor has null components. Now we are interested in the sources generating GWs, so the second part of Einstein equations is not zero. Combining Eq. 1.1 and Eq. 1.5 we can write:

$$\Box h_{ij} = -\frac{16\pi G}{c^4} T_{ij} \tag{1.8}$$

where we have omitted the signs ˜ and ' for simplicity.

In general, equations as

$$\Box f(\mathbf{x}, t) = s(\mathbf{x}, t) \tag{1.9}$$

where $f(\mathbf{x}, t)$ is the radiative field and $s(\mathbf{x}, t)$ is the source, can be solved by introducing the Green function $G(\mathbf{x}, t', \mathbf{x'}, t')$ defined as:

$$\Box G(\mathbf{x}, t', \mathbf{x'}, t') = \delta(t - t')\delta(\mathbf{x} - \mathbf{x'}) \tag{1.10}$$

and the expression of the radiative field $f(\mathbf{x},t)$ can be calculated solving the integral:

$$f(\mathbf{x},t) = \int dt' d^3x' G(\mathbf{x},t',\mathbf{x'},t') s(\mathbf{x'},t') \tag{1.11}$$

Using these informations, we can solve Eq. 1.8. The Green function associated to this equation is:

$$G(\mathbf{x},t',\mathbf{x'},t') = \frac{\delta(t' - [t - |\mathbf{x}-\mathbf{x}'|]/c)}{4\pi|\mathbf{x}-\mathbf{x}'|} \tag{1.12}$$

and the wave equation is solved by:

$$h_{ij}(t,\mathbf{x}) = 4\int d^3x' \frac{T_{ij}(t-|\mathbf{x}-\mathbf{x}'|,\mathbf{x}')}{|\mathbf{x}-\mathbf{x}'|} \tag{1.13}$$

We can approximate this formula for the case of great distances from the source. We define $r = |\mathbf{x}|$, so $|\mathbf{x}-\mathbf{x}'| = r - n^i x'_i + O(1/r)$ where $n_i = x_i/r$ and the previous expression becomes:

$$h_{ij}(t,\mathbf{x}) = \frac{4}{r}\int d^3x' T_{ij}\left(t - \frac{r}{c}, \mathbf{x}'\right) \tag{1.14}$$

This is the first term of the multipolar expansion. We show now that it is the quadrupolar order.

Using the properties of the energy-momentum tensor, in particular the identity: $\partial_i T^{ij} = 0$:

$$\left.\begin{array}{l} -\partial_t T^{tt} + \partial_i T^{it} = 0 \\ -\partial_t T^{ti} + \partial_j T^{ji} = 0 \end{array}\right\} \; \partial_t^2 T^{tt} = \partial_k \partial_l T^{kl} \tag{1.15}$$

and multiplying the last equation for $x^i x^j$, after some calculation, we obtain:

$$\partial_t^2(T^{tt}x^i x^j) = \partial_k\partial_l(T^{kl}x^i x^j) - 2\partial_k(T^{ik}x^j + T^{kj}x^i) + 2T^{ij} \tag{1.16}$$

which can be inserted in the Eq. 1.14 to obtain:

$$\begin{aligned} \frac{4}{r}\int d^3x' T_{ij} &= \frac{4}{r}\int d^3x' \begin{bmatrix} 1/2\partial_t^2(T^{tt}x'^i x'^j) + \\ \partial_k(T^{ik}x'^j + T^{kj}x'^l) + \\ -1/2\partial_k\partial_l(T^{kl}x'^i x'^j) \end{bmatrix} = \\ &= \frac{2}{r}\int d^3x \partial_t^2(T^{tt}x'^i x'^j) = \\ &= \frac{2}{r}\frac{\partial^2}{\partial t^2}\int d^3x' T^{tt}x'^i x'^j = \\ &= \frac{2}{r}\frac{\partial^2}{\partial t^2}\int d^3x' \rho x'^i x'^j = \\ &= \frac{2}{r}\frac{\partial^2}{\partial t^2} I^{ij} \end{aligned} \tag{1.17}$$

We have associated the wave tensor to the mass distribution of the source:

$$h_{ij}(t, \mathbf{x}) = \frac{2}{r} \frac{d^2 I_{ij}(t - r/c)}{dt^2} \tag{1.18}$$

Subtracting from the tensor I_{ij} its trace[4] I, we obtain the *quadrupole moment tensor*:

$$\mathbf{I}_{ij} = I_{ij} - \frac{1}{3}\delta_{ij} I \tag{1.19}$$

and if we introduce the tensor $P_{ij} = \delta_{ij} - n_i n_j$ (remember that $n_i = x_i/r$) we can write:

$$h_{ij}(t, \mathbf{x}) = \frac{2}{r} \frac{d^2 \mathbf{I}_{kl}(t - r/c)}{dt^2} P_{ik}(\mathbf{n}) P_{jl}(\mathbf{n}) \tag{1.20}$$

This means that Gravitational Waves are generated by mass motion of quadrupolar order.

However, the power P produced by the gravitational wave:

$$P = \frac{G}{5c^5} < \dddot{\mathbf{I}}_{ij} \dddot{\mathbf{I}}^{ij} > \tag{1.21}$$

is very low, being the factor $G/5c^5$ of the order of $10^{-53}\ W^{-1}$, so small that it is not possibile to produce gravitational waves in laboratories.

In the following paragraphs we list the most interesting sources of gravitational waves, focusing on the frequency band we are interested in this work.

1.2.1 Interesting source of Gravitational Waves

There are many theoretical models predicting transient gravitational wave emission in the few-kiloherzs range.

One of them is the gravitational collapse, including **core-collapse supernova** and long-soft gamma-ray burst scenarios, which are thought to emit gravitational waves in a frequency range above 1 kHz [8]. This process occurses when the iron core produced in the final stage of the nuclear burning exceeds the effective Chandrasekar limit [9, 10]. The object becomes gravitationally instable and the collapse happens, leading to dynamical compression of the inner core material to nuclear densities. According to the nuclear equation of state a rebound of the inner core takes place. A hydrodinamical shock wave propagates outward the core but the shock quickly loses energy, thanks to the dissociation of heavy nuclei and the production of neutrinos, so it stalls and must be revived to plow through the stellar envelope. The shock blows up the star and produces a Supernova explosion, behind a Neutron Star or a Black Hole.

[4]The trace of tensor A_{ij} is defined as $A = g^{ij} A_{ij}$

GW emission Process	**Potential explosion mechanism**		
	Magneto Rotational	**Neutrino**	**Acoustic**
Rotating collapse and Bounce	Strong	None/Weak	None/Weak
3d Rotational instabilities	Strong	None	None
Convection	None/Weak	Weak	Weak
Proto Neutron Stars g-modes	None/Weak	None/Weak	Strong

Table 1.1: *Overview on prominent GW emission processes in core-collapse Supernovae and their detection from Earth interferometers. For a galactic Supernova, 'strong' corresponds to probably detectable by initial and advanced LIGO [19], 'weak' means probably marginally detectable by advanced LIGO and 'none' means absent or probably not detectable by advanced LIGO.*

Iron core collapses are the most energetic astrophyical processes in the modern universe, liberating about 10^{53} erg of gravitational energy [8]. However, only 1% goes into the asymptotic energy of the Supernova explosion and becomes visible in the electromagnetic spectrum. It is not still clear which fraction of gravitational energy is transferred to revive the shock and ultimately unbind the stellar envelope. Three Supernova mechanisms are discussed in literature: the *neutrino mechanism* [10, 11], the *magneto-rotational mechanism* [12, 13] and the *acoustic mechanism* [14]. In iron core collapse and postbounce SN evolution, the emission of GWs is expected primarly from rotating collapse and bounce, non-axisymmetric rotational instabilities and proto neutron star pulsation. Additionally, anisotropic neutrino emission, global precollapse asymmetries in the iron core and surrounding burning shells, aspherical mass ejection, magnetic stresses and the late-time formation of black hole may contribute to the overal GW signature. While the Supernova rate in the Milky Way and the local group of galaxies is rather low and probably less than 1 Supernova per two decades [15], it could be 1 Supernova per year between 3-5 Mpc from Earth [16].

In particular, neutron star collapses resulting in rotating black holes emit GWs with a frequency greater than 2 kHz. From [17, 18] we can estimate the feasibility of these source detections. For an interferometric detector with the Virgo sensisitivity and for the signal coming only from the gravitational collapse of a rapidly and uniformly-rotating polytropic star at 10 kpc, we set an upper limit on the resulting signal-to-noise ratios of $SNR \approx 0.27 \div 2.1$ for Virgo/LIGO and $SNR \approx 1.2 \div 11$ for advanced LIGO [19].

Another potential class of high frequency gravitational wave sources is nonaxisymmetric hypermassive neutron stars resulting from *neutron star-neutron star mergers*. If the equation of state is sufficiently stiff, a hypermassive neutron star is formed as an intermediate step before a final collapse to a black hole, while a

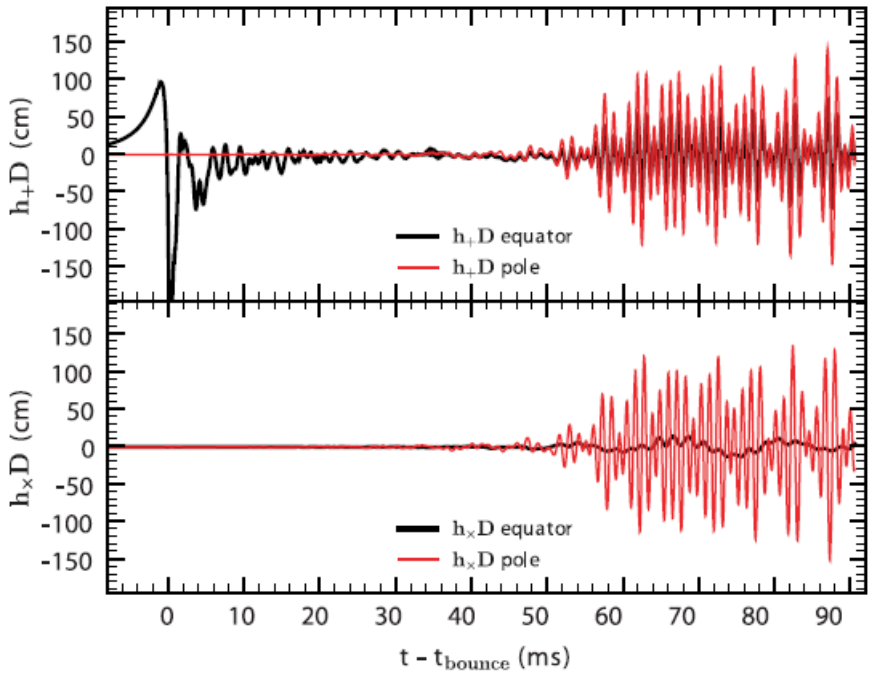

Figure 1.1: *h_+ (top) and $h_\times$ (bottom) multiplied by distance D as seen by observers along the equator (black) and along the pole (red) of the source for the model S20A2B4 from [20]. Note that the GW burst signal from core bounce is purely axisymmetric, since an axisymmetric system has vanishing $h_\times$ and vanishing GW emission along the axis of symmetry.*

softer equation of state leads to immediate formation of a black hole. Some models predict gravitational wave emission in the $2 \div 4$ kHz range from this intermediate hypermassive neutron star, but in many other cases higher frequency emission ($6 \div 7$ kHz) from a promptly formed black hole are foreseen [21, 22].

Other interesting sources of GWs are the **binary sistems**, composed by two astrophysical objects orbiting their mass center. Gravitational waves emission occurs during three phases: *inspiral*, *merger* and *ringdown*. In the inspiral phase, the two objects spiral towards one another and emit gravitational waves. GWs substract energy and angular momentum to the binary system and this loss causes a decrease of the rotation radius, with the two objects that start approaching each other. At a certain radius [21], they cannot fight the mutual gravitational force and they collapse and merge. The characteristics of the signal emitted during the collision are not exactly known, but numerical simulations calculate that the

emission period is very short (some millisecond). The last stage is the ringdown process, when the final object finds an equilibrium state, which is characterized by a series of oscillations of the object itself, the quasinormal modes.

During the inspiral phase, the two polarization components of the emitted wave depend on the various parameters involved in the motion.

Considering the following quantities:

- The Masses
  - The two object masses: M_1 and M_2
  - The total mass: $M_{tot} = M_1 + M_2$
  - The reduced mass: $\mu = \frac{M_1 M_2}{M_1 + M_2}$
  - The chirp mass: $M = \mu^{3/5} M_{tot}^{2/5}$
- The direction of wave propagation respect to the rotation plane
  - ι is the angle between propagation direction and rotation plane
- The distance R
- The coalesce time t_c

the two polarization components assume the values: [23]

$$\begin{aligned} h_+(t) &= \frac{4(GM)^{5/3}}{Rc^4} \frac{1+cos^2\iota}{2} \left(\pi f(t)\right)^{2/3} cos\psi(t) \\ h_\times(t) &= \frac{4(GM)^{5/3}}{Rc^4} cos\iota \left(\pi f(t)\right)^{2/3} sin\psi(t) \end{aligned} \tag{1.22}$$

where $f(t)$ is the rotation frequency and $\psi(t)$ is the wave phase, defined as:

$$f(t) = \frac{1}{\pi} \left(\frac{256}{5} \frac{(GM)^{5/3}}{c^5} (t_c - t) \right)^{-3/8} \tag{1.23}$$

$$\psi(t) = -2 \left(\frac{G^{5/3}}{c^5} \right)^{-3/8} \left(\frac{t_c - t}{5M} \right)^{5/8} + cost \tag{1.24}$$

It is not easy to obtain an equation of the merge process, however numerical simulations can describe this process and the corresponding wave emission.

The Ringdown process, instead, is well characterized. The wave amplitude is expressed by the following relation: [24, 25]

$$h(t, \iota, \psi_0) = \frac{A(\epsilon, f_0, Q)}{r} e^{-\frac{\pi f_0}{Q} t} cos(2\pi f_0 t - \psi_0) \tag{1.25}$$

where we have defined:

$$\begin{gathered} f_0 = \frac{c^3}{2\pi GM}[0.63(1-a)^{3/10}] \\ Q \approx 2(1-a)^{-9/20} \\ a = S\frac{c}{GM^2} \end{gathered} \tag{1.26}$$

The detectability of such processes are calculated in [21] for a distance of 10 kpc and an energy efficiency in emission of gravitational raditation equal to $E \approx 10^{-7} \div 10^{-6}$. The expected SNR assumes a value of 5 for interferometer with the same sensitivity as LIGO. Considering a minimal value of 3 for a detection in case the preceding inspiral chirp has been measured, such GW signals may be identified up to 35 Mpc. According to [26, 27] these distances correspond to event rates of $0.004 \div 0.5$ per year.

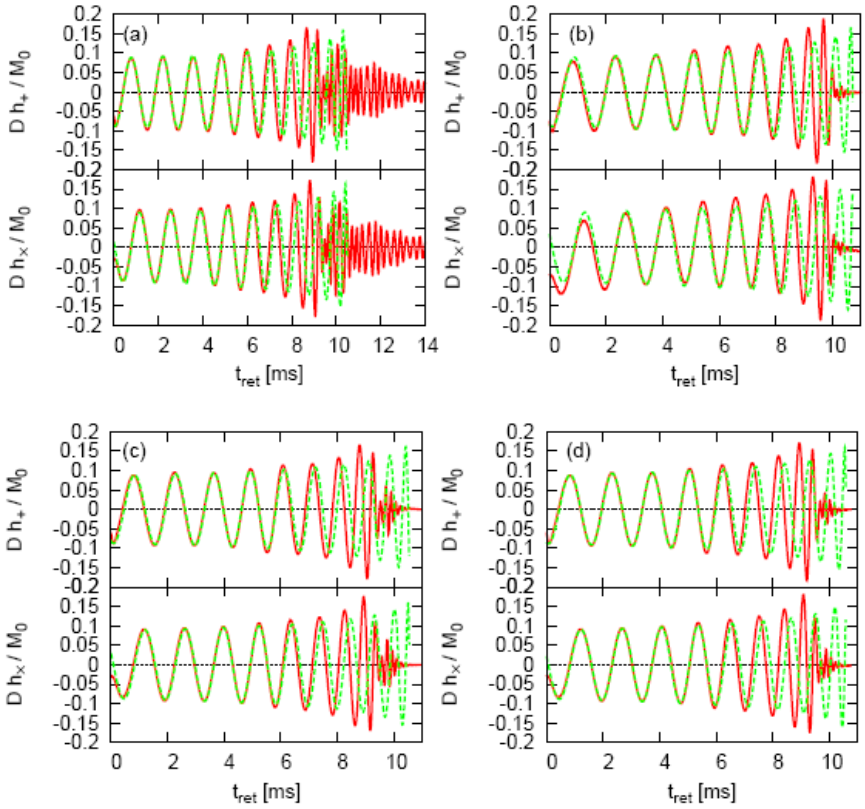

Figure 1.2: *h_+ (top) and $x_\times$ (bottom) for several models of binary merging ((a) APR1414H, (b) APR1515H, (c) APR1316H, (d) APR134165H, see [22] for details). Solid and dashed curves denote respectively the waveforms calculated by the simulation and a formula using Taylor expansion.*

An interesting similar scenario comes from the merging of low-mass black holes

[28]. The typical frequency of the emitted gravitational wave can be associated to the total mass M and to the orbital inspiralling period T for small periods:

$$f_g \approx 3 \cdot 10^3 \left(\frac{M}{10^{-7} M_\odot}\right)^{-5/8} \left(\frac{T}{5yr}\right) \quad \text{Hz} \tag{1.27}$$

The strengh of gravitational waves from a monochromatic source observed by using an interferometer is evaluated as:

$$h \approx 1 \cdot 10^{-24} \left(\frac{M}{10^{-7} M_\odot}\right)^{274/333} \left(\frac{f_g}{100Hz}\right)^{-2/9} \left(\frac{\Delta T}{5yr}\right) \tag{1.28}$$

where the above estimate is valid only for the frequencies below f_g for which the source frequency does not change rapidly during the observation time $\Delta T << T$ [28].

The merging of a binary system is related also to the **Gamma Ray Bursts** (**GRB**) which are of great interest for the GW physics. GRBs are short intense flashes of soft ($\sim$ MeV) γ-rays. The observed bursts arrive from apparently random directions in the sky and they last between tens of milliseconds and thousand of seconds. GRB duration follows a bimodal distribution with a minimum around two seconds. Oservations have shown that bursts with duration shorter than this minimum are composed, on average, of higher energy photons than longer bursts. There two populations (usually labelled with *short* and *long*) are related to different physical phenomena. Long GRB are produced by a catastrophic event involving a stellar-mass object or system that releases a wast amount of energy ($\gtrsim 10^{-3} M_\odot s$) in a compact region ($< 100$ km) on time scales of seconds to minutes [29, 30, 31]. Regarding the short GRB, the leading progenitor candidate is the coalescence of a neutron star with another neutron star or with a black hole [30, 32, 33, 34, 35]. Since the GRB source are related to the emission of GW, GRBs are considered of great interest by the GW community, being electromagnetic counterparts of the most accessible GW sources for ground-based GW observatories. Short GRBs may emit GWs also if the progenitor is not a compact binary merger, like a collapse of a rotating compact object to a black hole. However, the amplitude of these waves is highly uncertain [36].

Other possible source of few-kilohertz gravitational wave emission include *neutron star normal modes* (in particular f-mode) [37] as well as neutron stars undergoing torque-free precession as a result of *accreting matter* from a binary companion [38] or some scenarios for gravitational emission from *cosmic string cusps* [39].

Chapter 2

Gravitational Waves Detectors

GW detectors aim to measure the distance variation caused by the passing of a gravitational wave. The first who tried to detect gravitational waves was Joseph Weber who in 1957 developed the first prototype of the acoustic detectors. Acoustic detectors measure the oscillations of a bar which resonates at the passage of a gravitational wave. Acoustic vibrations are converted in electronic signals and so measured. The bar must be seismically and thermically isolated to decrease the effect of thermal and enviromental noise.
Acoustic detectors present limited sensitivity in a very narrow frequency band of detection and it is very difficult to extend the frequency range of sensitivity. AURIGA [40], the most sensitive acoustic detector at the present time, has a frequency range of about 100 Hz near the 1 kHz with a sensitivity of about $2 \cdot 10^{-21}\mathrm{Hz}^{-1/2}$. It is therefore supposed to detect GWs from strong astrophysical sources in the Milky-Way galaxy, such as supernova collapses, see [41].

The GW community has developed a different type of detectors, the interferometric ones. Such detectors use the principle of the light interference of Michelson-Morley experiment. A laser beam is splitted in two components which follow two perpendicolar arms. If a gravitational wave interacts with the detector, the two arms will modify their length, and the two components of the beam will create a not null interference figure. The measure of this interference can be connected to the effect of the gravitational wave. Because of the small values involved ($\Delta L/L \approx 10^{-21}$) the interferometer arms are constructed as long as possible with respect to the Earth curvature and to the construction issues and they reach the length of few kilometers. Some examples of such detectors are Virgo (in Italy) [42], the three LSC located in the USA [43], GEO (in Germany) [44] and TAMA [45] in Japan. GW community is planning to construct another earth detector, AIGO [46] in Australia. Additionally, a space detector is under study, LISA [47] that will consist of three satellites marking the vertices of an equilater triangle and orbiting around the sun. Hovewer, the LISA detector is designer to study the very

low frequency: $10^{-4} \div 10^{-1}$ Hz.

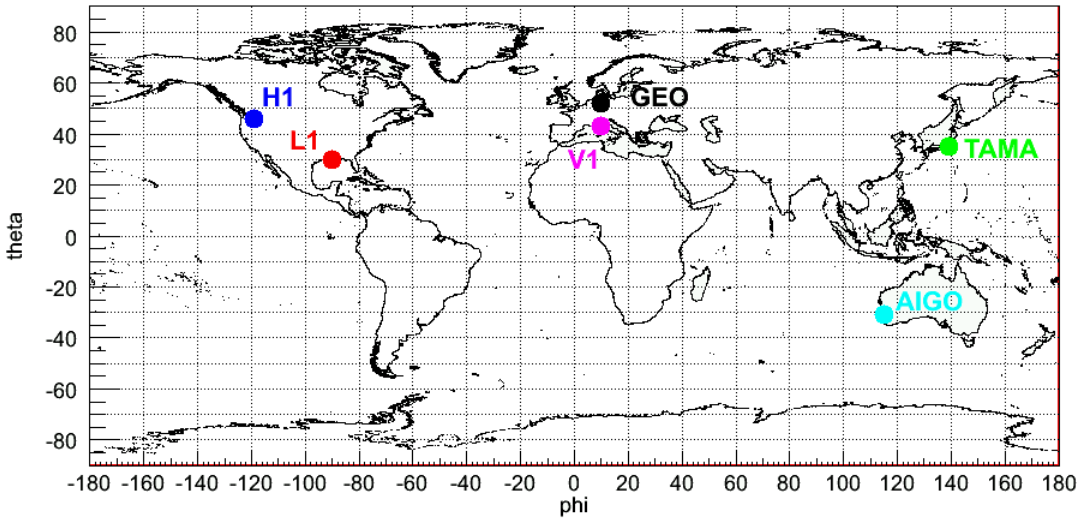

Figure 2.1: *Location of the working earth interferometers: LSC Livingston (L1) and Hanford (H1), Virgo (V1), GEO600 (GEO) and TAMA. Established location of the projected detector AIGO.*

2.1 Virgo and the interferometric detectors

For this work we considered data of the Virgo and LSC detectors.

Virgo (**V1**) is an interferometric detector located in Cascina (Pisa, Italy), constructed by an Italian-French collaboration. Its arms are of 3 km length.

One of the LSC interferometers (**L1**) is located at Livingstone (Lousiana) and the other two (**H1** and **H2**) are co-located in the same site at Hanford (Washington State). L1 and H1 arms are 4 km, while H2 arms are only 2 km long. H2 has not been working during the last period (mid-2009).

The different detectors have similar characteristics, such as: the beam-splitter, the reflecting mirrors and vacuum tubes for the laser beam. The arms contain Fabry-Perot cavities which increase the beam optic length. The reflecting mirrors are smoothed with an high precision and their absortion and diffusion are very small. The laser power and environmental noises produce disturbing motion of the mirrors that are attenuated by seismic isolators. These detectors present sensitivity less than $10^{-20}\ Hz^{-1/2}$ up to about 40 Hz.

We describe the main characteristic of an interferometric detector focusing on Virgo properties.

Virgo optical scheme is composed of a laser with power equal to 20W emitting at 1064 nm. La laser radiation is phase modulated at 6 and 14 MHz before entering in the void system through the injection bank. The light is stabilized with a mode-cleaner using the Pound Drever Hall method [48, 49]. A stabilization at low frequency is made to lock the laser frequency on the length of 30 cm on the reference triangular monolithical cavity (RFC), which is suspendend in the void. The laser light that is emitted by the injection system enters in the interferometer (ITF) passing through the recycling mirror (PR). This mirror reflects the light incoming from the interferometer, so it increases the circulating light.

Then the laser passes through the Beam Splitter, a particular mirror which splits the laser in two components that enter in the two orthogonal Fabry-Perot cavities. The Fabry Perot cavities have a length of 3 km and an optical recycling gain equal to 50. At the end, the circulating power in the interferometer is about 16kW.

The gravitational signal is extracted by the detection bench (DB) which works in normal pression conditions. From the principal beam detected by the DB, other minor beams are extracted for control purpouses.

The Virgo interferometer is a *null-instrument*, i.e. it works in the *dark fringe*. This means that when the signal is absent, the detector gives no signal, otherwise the signal is not zero when the GW is present. This is made possible using the detection technic called Pound-Dreiver, that gives in exit an electrical signal, which is linearly coupled with other signals when the GW is present.

Virgo has implemented a particular attenuation system: the super-attenuator (SA). This is a passive attenuation system composed of a chain of five attenution phases and an initial one: the superfilter. The last has the same characteristics of the others, and it is supported by an inverted pendulum with three legs. The working principle is that for a pendulum when the force frequency f is far from the resonance frequency f_0 the transfer function is proportional to $(f_0/f)^2$. For a chain of N pendulums the attenuation of seismic noise on the suspension point with respect to the mirror is proportional to $(f_0/f)^{2N}$.

However, the horizontal and vertical components of the seismic noise are comparable at Cascina, so an attenuation on the vertical motion is necessary, to avoid coupling between vertical and horizontal motions of the mirrors. This is realized by the combined effect of a system of metallic blades and magnetics anti-springs that allows to reach the desidered sensibility [50].

The coupling between horizontal and rotational degrees of freedom is reduced by the large inertia momentum of each filter, and by reducing the arms of the turning momentum on each filter. The SA has been projected with five pendulums

to reach the desidered sensitivity at 10 Hz. The transfer function is well attenuated at high frequencies, but not at lower frequencies. The superfilter is projected to reduce the resonance frequency. The three legs with length of 6 m are substained on a ring with elastic joints. A shift on the superfilter is related to a small deformation on the joints. So, the elastic force on the joints is very small, and this decreases the resonance frequency of the system.

The test masses should be continously mantained on the working point, together to the interferometer. This is realised with a damping active procedure by a control system regulated by accelerometers which measure the system acceleration on vertical, horizontal and rotational movements. These accelerometers are posed on the superfilter and on the ring on which are the three legs. This system composed by three legs and superfilter is an optimal system which induces big movements on the suspension point in a soft way, i.e. low frequency motions using a small force on the joints. Moreover, the ring is mantained continuosly horizontal with the use of motors.

The last phase is composed by the *marionetta*, which is connected by cables of length 1 m. With the marionetta, the test mass is posizioned on the 6 degress of freedom with movements on the order of the laser length wave. In this way it is possible to compensate the seismic residual noise on the SA chain. These shifts are made using electromagnetical attenuators located on the last phase of the SA or on a reference mass suspended on the marionetta.

2.2 Noise sources for an interferometer detector

An instrument with such a complex structure is obviously affected by many noise sources. In this chapter we treat the main sources of noise for interferometers detectors. We suppose that the stochastic process associated to noise sources is gaussian and stationary, so we can study the processes with approximations of low order (gaussianity) and describe them in the frequency domain (stationariety). We can therefore use the noise spectral density to describe the noise contributes to the total sensitivity. If $h(t)$ is the temporal series of the detector, we define the sensitivity as:

$$h(f) = \sqrt{2S_h(f)} \tag{2.1}$$

where $S_h(f)$ is the two-sided spectrum[1], defined as:

$$S_h(f) = \lim_{T\to\infty} \left| \frac{1}{T} \int_{-T/2}^{T/2} h(t) e^{2i\pi ft} dt \right|^2 \tag{2.2}$$

[1]The 2 factor is introduced because we consider only the spectrum with positive frequencies (one-sided spectrum).

The behaviour of the Virgo sensitivity curve as a function of frequency is reported in the Fig. 2.2.

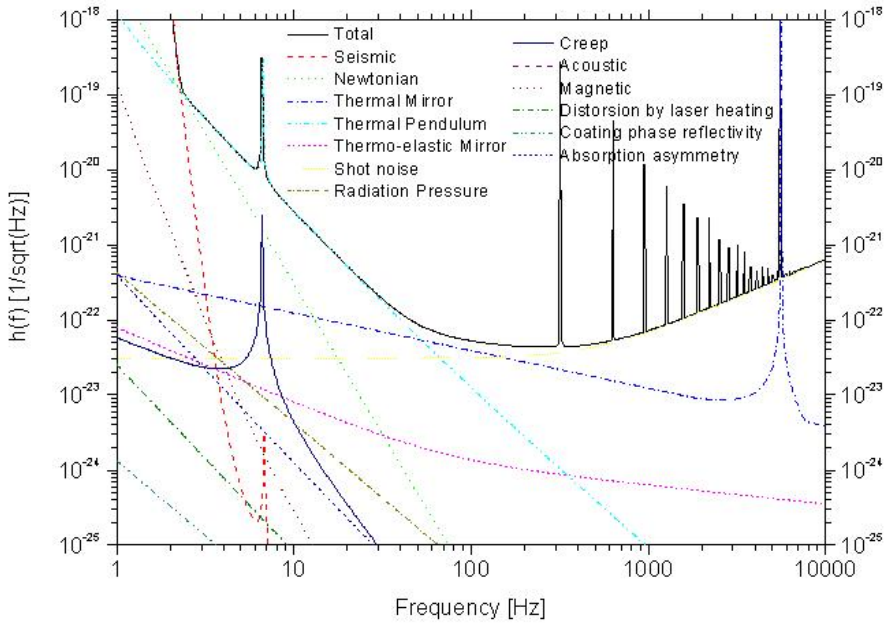

Figure 2.2: *Virgo design sensitivity curve including the noise sources*

Here we list the main contributions of the noise sources for an interferometer detector.

2.2.1 Seismic noise

This noise dominates at low frequency and it is due to human and geology activity. In the $1 \div 10$ Hz band this noise is primarly produced by winds, local traffic, train passage, etc...[51]. Atmosferic cyclones produce fluctuations at lower frequencies. The spectral amplitude of the seismic noise has two main peaks: on the oceanic wave period (12s) and on double frequency, while in the $1 \div 10$ Hz can be approximed by a decreasing power law.

In the Cascina site the seismic noise can be parametrized with the spectrum:

$$|x_{seismic}(f)| = \begin{cases} \frac{10^{-10}}{f^3} & f < 0.05 \text{ Hz} \\ 10^{-6} & 0.05 < f < 0.3 \text{ Hz} \\ \frac{10^{-7}}{f^2} & f < 0.3 \text{ Hz} \end{cases} \tag{2.3}$$

The seismic noise is attenuated by the pendulum chain which holds the test masses. Taking in account the superattenuator transfer function (TF), the sensitivity can be expressed by:

$$h(f) = \frac{2}{L} TF \sigma_{h-v} |x_{seismic}(f)| \tag{2.4}$$

where $\sigma_{h-v} = 10^{-2}$ is the coupling constant between superattenuator horizzontal and vertical motions and L is the interferometer arms length.

2.2.2 Shot noise

The optical length difference between the light path along the two arms is measured with the optical power variation at the end of interferometer. Thus, the sensitivity of GW detection depends on the minimal power variation we are able to measure, i.e. the precision of the counted number of photons that arrive to the photodiode.

The distribution of photon follows the Poisson statystics that can be approximed for large number of photons to a Gaussian distribution with standard deviation $\sigma = \sqrt{N}$ if N is the mean number of photon. If we consider as a working point the laser power on the exit P_{out}, which is a half of the laser power on entrance P_{in}, the mean number of photon can be expressed as:

$$N = \frac{\lambda}{2\pi\hbar c} \frac{P_{in}}{2} \tau \tag{2.5}$$

where λ is the wave length of the laser, and τ the measure time. The fluctuation on the differences of test masses position is:

$$\sigma_{\delta L} = \frac{\sigma}{N} / \frac{1}{P_{out}} \frac{dP_{out}}{dL} = \sqrt{\frac{\hbar c \lambda}{4\pi P_{in} \tau}} \tag{2.6}$$

and from this equation, considering that $P_{out}/dL = 2\pi\lambda P_{in}$ we can write the relative fluctuation in sensitivity:

$$\sigma_h = \frac{\sigma_{\delta L}}{L_e} = \frac{1}{L_e} \sqrt{\frac{\hbar c \lambda}{4\pi P_{in} \tau}} \tag{2.7}$$

where L_e is the effective optical length of the interferometer arms.

Since no characteristic scale of this noise is present, the spectral density is constant with amplitude:

$$h(f) = \frac{1}{L_e} \sqrt{\frac{\hbar c \lambda}{4\pi P_{in} \tau}} \tag{2.8}$$

for a two-side spectrum. For high frequencies, i.e. small measured times, this approximation is not more valid because the photon number is too small.

From equation 2.7 we can see that we can improve the sensitivity increasing the optical length. This is why the Virgo team has introduced the Fabry-Perot cavities [52], we have in fact that:

$$L_e = \frac{2}{\pi} FL \tag{2.9}$$

where F is the cavity finesse and L the physical arms length. However, the effective length should not be too long, because this would increase the storage time in the cavity. If the storage time is comparable to the GW period, the cavity means the GW received signal on more periods. For high frequencies the interferometer sensitivity descreases and the shot noise can be written as:

$$h(f) = \frac{\pi}{2FL}\sqrt{\frac{\hbar c\lambda}{4\pi P_{in}\tau}}\sqrt{1+\left(\frac{f}{f_c}\right)^2} \tag{2.10}$$

where $f_c \cong 500$ Hz.

2.2.3 Radiation Pressure noise

The shot noise can be reduced by increasing the laser power. However, the light produces a force effect on the mirrors which is proportional to the laser power ($F_{rad} = P/c$). This force produces a mirror fluctuation:

$$h(f) = \frac{1}{m(2\pi f)^2} F(f) \tag{2.11}$$

where the transfer function is considered far from the resonance frequency. The fluctuation on the two arms are uncorrelated, so the spectral density amplitude related to this noise is given by:

$$h_{rad}(f)\frac{1}{mf^2L}\sqrt{\frac{\hbar P_{in}}{2\pi^3 c\lambda}} \tag{2.12}$$

2.2.4 Quantum noise

Radiation pressure noise dominates at low frequencies, while for high frequencies the shot noise dominates. When these two contributes are equal (at a certain frequency) the total spectral density assumes a minimum value, sum of the two contributes. This is verified if we choose the power in entrance:

$$P_{in}(f) = \pi c\lambda m f^2 \tag{2.13}$$

In this case we say that the quantum limit is reached:

$$h(f) = \sqrt{\frac{1}{\pi f L}}\sqrt{\frac{\hbar}{m}} \tag{2.14}$$

This is not a true pover density, but only a limit curve which could be reached only at a certain frequency if we conveniently choose the laser power.

2.2.5 Thermal Noise

The detector components work at a environmental temperature, so they are influenced by the thermal noise, related to the dissipating part of the system. This relation can be expressed by the **Fluctuation and Dissipation** theorem [53], if we suppose that the system is linear and in thermodinamic equilibrium.

If the system is subjected to an external force $F_{ext}(f)$ it will move at the velocity $v(f)$. Because of the linearity of the system, we can write:

$$F_{ext}(f) = Zv(f) \tag{2.15}$$

where Z is the system impedance.

The fluctuation-dissipation theorem says that the power spectrum $F^2_{thermal}(f)$ of the acting force is given by:

$$F^2_{thermal}(f) = 4K_B T Re(Z(f)) \tag{2.16}$$

where t is the system temperature, k_B the Boltzmann constant and $Re(Z(f))$ is the impedance real part (i.e. the dissipative part). This theorem allows to obtain an estimation of the thermal noise without knowing the microscopic processes that cause the dissipation, but only introducing a macroscopic model of impedance as a function of frequency.

Using this theorem we can calculate the thermal noise due to the mirror movements [54], that are due to three main factors:

- pendulum motion: x_p
- normal modes of the test masses: x_s
- violin mode: x_v

The total contribution to the sensitivity is given from:

$$h_{thermal}(f) = \frac{1}{L}\sqrt{x_p(f)^2 + x_s(f)^2 + x_v(f)^2} \tag{2.17}$$

An example of the sensitivity curves of LIGO-Virgo detectors during S5/VSR1 is reported in Fig. 2.3

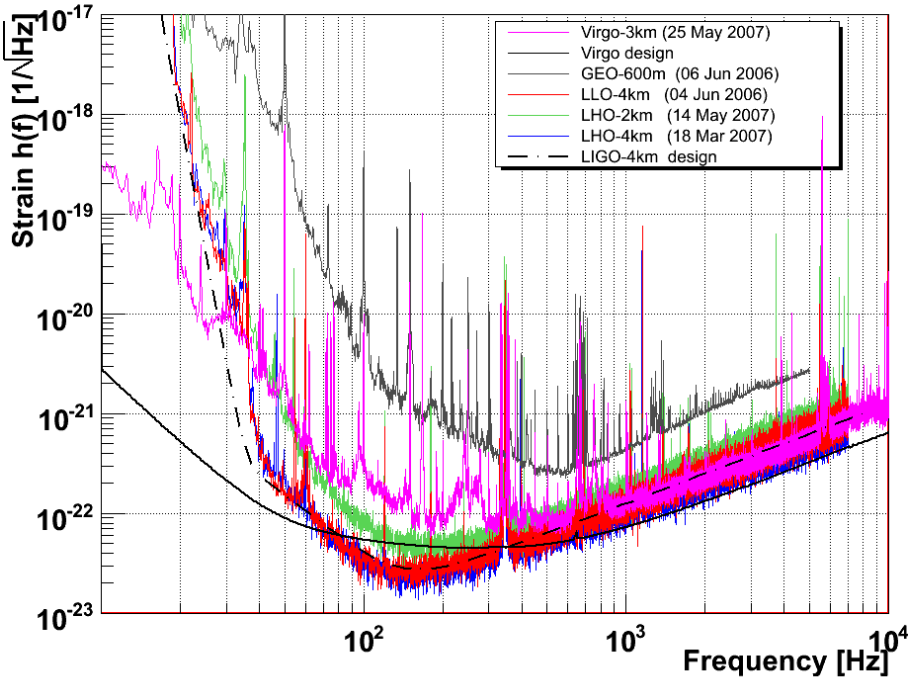

Figure 2.3: *V1, L1, H1 and H2 real sensitivity curves during S5-VSR1 run.*

2.3 Detector response

The response of a single detector $X(t)$ is the sum of the noise $n(t)$ and of the eventual contribution due to the GW interaction $\xi(t)$: $X(t) = \xi(t) + n(t)$. For instance, for an interferometer with equal arms of length L:

$$\xi(t) = [\delta L_x(t) - \delta L_y]/L \tag{2.18}$$

where $\delta L_{x,y}$ are the displacements produced by the GW. A very useful mathematical tool is the *detector tensor* D_{ij} which allows to express the relation between the scalar output $\xi(t)$ and the GW tensor $h_{ij}(t)$:

$$\xi = D^{ij} h_{ij}(t) \tag{2.19}$$

It is obvious that in the TT gauge the detector tensor assumes a simple form, and so the response can be written as:

$$\xi(t) = F_+ h_+(t) + F_\times h_\times(t) \tag{2.20}$$

where F_+, $F_\times$ are called **Antenna Pattern** and depend only on the relation between the source direction and the interferometer orientation. If we consider the Euler Angles defined in Fig. 2.4 between the wave frame and the detector arms (which have directions of the x,y-axis), the antenna patterns assume the form: [55]

$$\begin{cases} F_+(\theta,\phi,\psi) = \frac{1}{2}(1+cos^2\theta)cos2\phi cos2\psi - cos\theta sin2\psi sin2\phi \\ F_\times(\theta,\phi,\psi) = \frac{1}{2}(1+cos^2\theta)cos2\phi sin2\psi - cos\theta sin2\psi cos2\phi \end{cases} \tag{2.21}$$

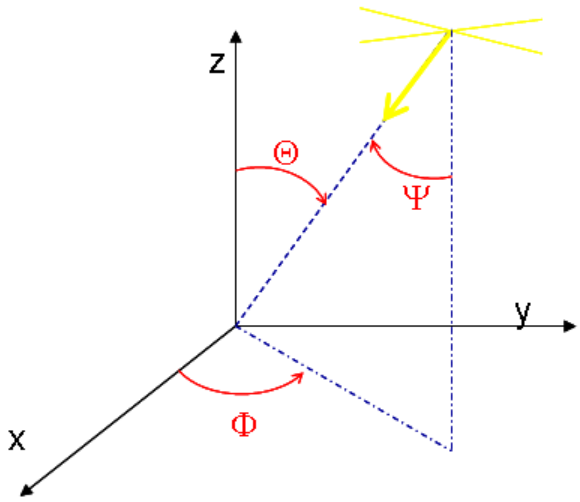

Figure 2.4: *Relationship between TT wave frame and detector frame. For an interferometer, the arms are located at the x,y-axis. The angle (θ,ϕ,ψ) are the Euler Angle of the transformation between detector frame and wave frame.*

The Fig. 2.5 shows the $\sqrt{(F_+^2 + F_\times^2)}$ quantity for the single detectors. The values of antenna patterns vary according to the points on the earth were they are calculated. The sensitivity of the single detector depends on the arms orientation: the directions of the arms are the less sensitive, where as the more sensitive are the orientations orthogonal to the plane defined by the two arms. This can be easily seen on the Fig. 2.5, the more sensitive zones (greater z-axis values) are near the earth position of the detectors.

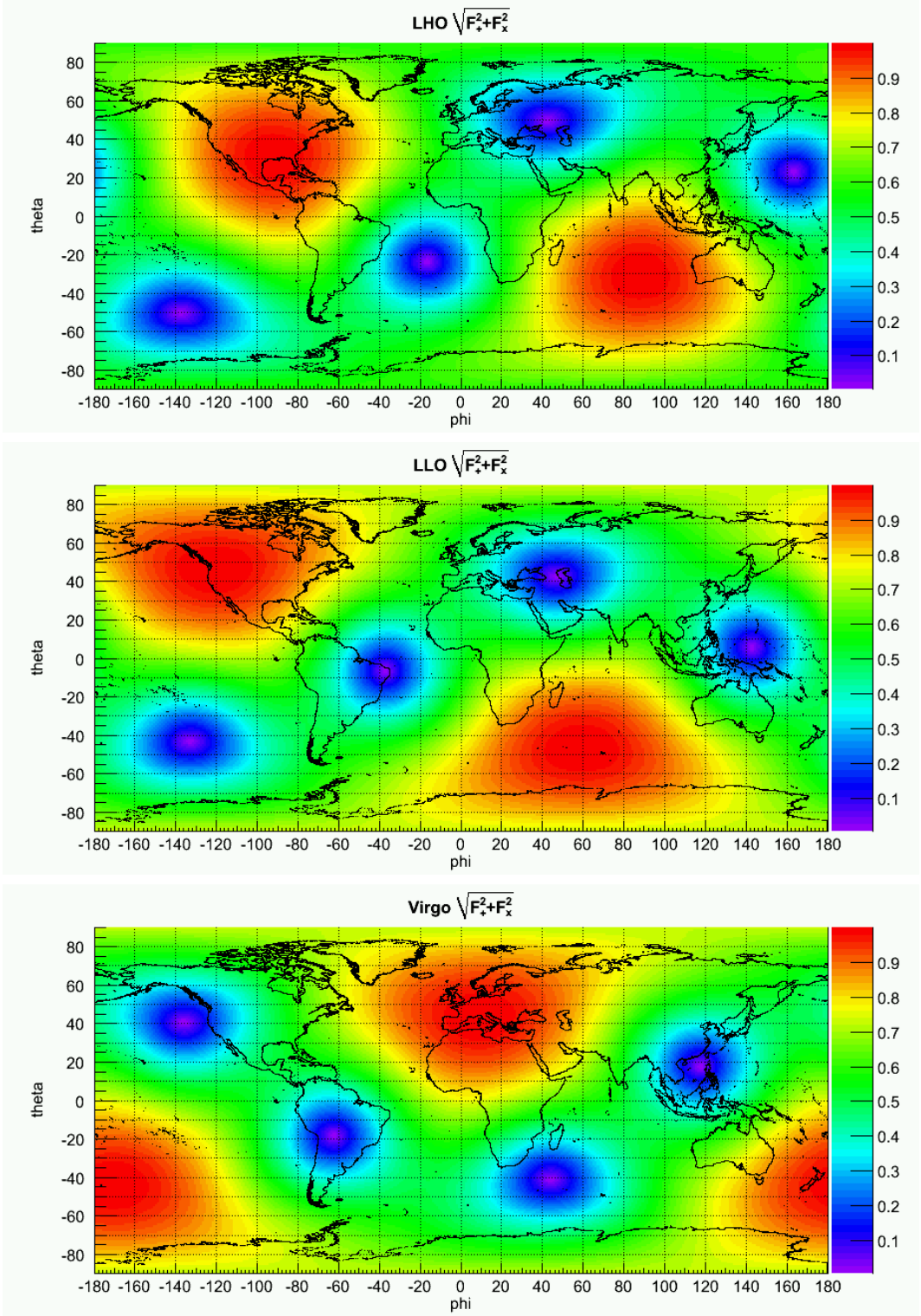

Figure 2.5: *Variation of* $\sqrt{F_+^2 + F_\times^2}$ *quantity following earth coordinate. On x-axis is reported the longitude, and on y-axis the latitude. Coloured axis reports valued of antenna patterns. From top to bottom the detectors are: L1, H1, V1.*

Chapter 3

Waveburst: an algorithm of Burst Search

With the word **Burst** we indicate signals with short duration, i.e. less than 1 second. This collection includes a large variety of possible waveforms, which could be wide or narrow banded in the frequency domain. We have just explained how Bursts GWs could be generated by complex astrophysical processes, such as Supernovae collapse or merging of binary systems. Hence, in general, analysis method searching for Bursts must be robust with respect to the expected waveforms. This means that we should be ready to the unexpected, because we do not know exactly all the GW types that could be detected from the detectors. We can relax the algorithm robustness in case some informations about the source are known thanks to other astrophysicals observation. However, it is convenient that the algorithms remain as robust as possibile against the possible variety of waveforms even if we have accurate informations on the espected GW, to keep the algorithm as general as possible.

According to the main expected sources of GW Bursts, there are principally two ways of bursts searching. The first does no make assumptions about the signal (for instance: direction source, waveform type, time of arrival...), the second takes in account the informations from an external trigger (in this case source location and timing are known). For both the cases, the common parts are the main step of the analysis: from the generation of triggers to the estimation of efficiency and accidental background.

The trigger generator algorithms (or analysis pipelines) searches at local excesses of energy in the calibrated h(t) time series [56, 57, 58, 59]. Up to now, the algorithms involve data from sets of detectors (**networks**), since detections from single detector cannot be as realiable as those from multiple detectors. This is due to different factors: one is that no GWs have never been detected, so we have no possibility to comparisons. Another is the difficulty to extract a possible

GW signal from the various noise sources characteristics of each detectors Sec. 2.2. For these reasons using a single detector it is not so easy to understand if a noise excess is due to a GW or to an enviromental disturbance. But if this noise excess is seen contemporaneously from more detectors, it could be due to a certain correlation within the different detectors. Otherwise, this correlation could also be casual. There are two ways to combine network data: incoherently and coherently. The **incoherent** pipelines generates the triggers singularly for each detectors, where as the **coherent** ones generate an unique list of triggers combining data from several detectors. In the first case, various decision rules (time coincidences, time-frequency coincidences, cross-correlation data stream, ...) select coincident triggers from all detectors and consider them as network events. Semi-coherent approaches have been adopted, starting with a separate generation of triggers and using a data coherent combination from such triggers. An example of semi-coherent approach can be found in [60].

Coherent methods allow to implement more consistency relationship on the data stream between detector signals, so it is possible to reconstruct the burst waveform and the source location in a more reliable way with respect to incoherent pipelines. However, also with incoherent methods it is possible to identify the source coordinates using a network composed of almost three detectors, with a certain error which depends on algorithm method and on data noise. Recent studies show that the typical angular accuracy is of the order of a few degrees, depending of the GW direction with respect to the detector plane [61, 62].

The search of GW Bursts is based on the estimation of a detection confidence for a definite false alarm. This is implemented characterising the detection efficiency and the background estimation. To get the efficiency, "fake" waveforms are injected in the data stream adding the wave amplitude to the noise data; then the algorithm is applied to count how many injected signals are detected. Generally, for these tests, the predicted astrophysical waveforms are substituted by generic signals such as Gaussian or SineGaussian pulses. These kinds of signals are interesting because they span the detector bandwidth and provide a good sensitivity estimate for the whole range of accessible frequencies. An useful quantity to characterize the injected signal is the so called "h_{rss}", which corresponds to the total GW signal energy:

$$h_{rss} = \sqrt{\int_{-\infty}^{+\infty} h(t)^2 dt} \tag{3.1}$$

in units of $Hz^{-1/2}$. We can compare it directly to the detector sensitivities (amplitude spectral densities), expecially when the injected waveform is well localized in frequency.

The background is estimated by shifting one or more detectors data with re-

spect to the others and processing these time shifted data sets with the search algorithms. The relative time shifts are chosen to be much larger than the light times of flight between the detectors and than the expected signal durations, to be sure that reconstructed events cannot be due to real GW passage, but they are false alarms, or *glitches*. We can select events by making use of decision rules to obtain a definite false alarm rate, and apply these rules to the not shifted stream (*zero lag*); the final list is the list of candidate events at a confidence level given by the previously chosen false alarm rate.

For the triggered search, this procedure is quite different, because the data are splitted into an on-source region where the GW is expected and on a background region where the noise is expected to be statistically similar to the one in the on-source region. The background region generally consists of a few hours from each side on the on-source region. All the efficieny and background studies are performed with the background data set.

The rejection of loud glitches is performed applying data quality flags and event by event vetoes. The first indicate when one interferometer is not working properly during a certain period of time. The seconds are imposed when an excess of coincidence is found between the gravitational strain channel triggers and auxiliary (enviromental) channel events. Data quality and event by event vetoes are not generated by the search pipeline, but by different algorithms, and allow to disregard short time intervals, saving therefore the rest of the detector observation time.

The LSC-Virgo collaboration has implemented various algorithms of Burst Search. In this chapter we explain the **Waveburst** pipeline, an algorithm of data analysis developed into the LSC collaboration that does not make a priori assumptions on the signal. The algorithm is targeted to short signals, i.e. with duration less than one second. This method has been first designed as an **incoherent** algorithm [56, 63], with the possibility to apply it to a single detector or to make time-frequency coincidence between more than one; then it has been modified to implement a **coherent** analysis on a network of detectors using a likelihood method [57, 64].

The pipeline consistes of two stages: the production stage, when the burst triggers are generated, and the post-production stage, when addictional selection cuts are applied to distinguish GW candidates from the background events.

3.1 Likelihood method

Coherent Waveburst (**CWB**) introduces a coherent approach to combine data stream from different detectors, based on the **Constrained Likelihood** method to combine **coherently** the detector data stream [57, 64].

It is convenient to introduce a useful formalism to calculate the likelihood

definition. We recall the definition of the detector response to a gravitational signal ξ:

$$\xi = F_+ h_+ + F_\times h_\times \tag{3.2}$$

To simplify notations, we can define complex waveforms and antenna patterns as:

$$\zeta = h_+ + i h_\times \tag{3.3}$$

$$A = F_+ + i F_\times \tag{3.4}$$

where i is the imaginary unity. With such notation, the detector response becomes:

$$\xi = \zeta \cdot \tilde{A} + \tilde{\zeta} \cdot A \tag{3.5}$$

and a coordinate transformation in the wave coordinate frame is performed by a rotation of the type:

$$\begin{aligned} A' &= e^{i\psi} A \\ \zeta' &= e^{i\psi} \zeta \end{aligned} \tag{3.6}$$

hence, a coordinate transformation does not change the detector response.

3.1.1 A single detector

We start calculating the Likelihood for a single detector, such as we can apply the hypothesis test. This test requires the definition of a *decision rule* to select one of the two mutually exclusive hypotheses: the absence (H_0, null) and the presence (H_1, alternative) of the signal in the data stream.
Considering $x = \{x[1], x[2], ..., x[I]\}$ as the detector data, the H_0 and H_1 are described by the two probability densities $p(x|H_0)$ and $p(x|H_1)$. Any decision rule is associated to a threshold applied to these densities, and it is characterized by the table below.

		Signal presence	
		Yes	No
Hypothesis	Yes	True Alarm	False Alarm
	No	False Dismissal	True Dismissal

Table 3.1: *Right and wrong possible outcomes from decision rules. Each outcome has an associated probability.*

The Neyman-Pearson criterion says that when H_1 is a simple hypothesis, the optimal decision rule has the lowest false dismissal probability for fixed false alarm

probability, i.e. the rule rejects H_0 when the *likelihood ratio* $\Lambda(x)$ is greater than a threshold value fixed by the specified false alarm probability. $\Lambda(x)$ is defined as:[1]

$$\Lambda(x) = \frac{p(x|H_1)}{p(x|H_0)} \tag{3.7}$$

In the GW data analysis, under the assumption of Gaussian white noise with zero mean, the probabilities densities are:

$$p(x|H_0) = \prod_{i=1}^{I} \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{x^2[i]}{2\sigma^2}\right) \tag{3.8}$$

$$p(x|H_1) = \prod_{i=1}^{I} \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{(x[i]-\xi[i])^2}{2\sigma^2}\right) \tag{3.9}$$

where σ is the standard deviation of the noise.

We use the logaritmic value of the likelihood ratio, which we call *likelihood functional* (o simply likelihood):

$$L = ln(\Lambda(x)) = \sum_{i=1}^{I} \frac{1}{\sigma^2}\left(x[i]\xi[i] - \frac{1}{2}\xi^2[i]\right) \tag{3.10}$$

3.1.2 A network of detectors

There are no theorems that extend the Neymann-Pearson criterion from one to more detectors. However, we apply the same formalism for a network of N detectors as a starting point for the likelihood approach. Detector response of the k^{th} detector is given by:

$$\xi_k = \zeta_k \cdot \tilde{A}_k + \tilde{\zeta_k} \cdot A_k \tag{3.11}$$

and the total likelihood becomes:

$$L = \sum_{k=1}^{N}\sum_{i=1}^{I} \frac{1}{\sigma_k^2}\left(x_k[i]\xi_k[i] - \frac{1}{2}\xi_k^2[i]\right) \tag{3.12}$$

where we suppose that detector noises are not correlated with each other.

[1] In the case that H_1 is a multiple hypothesis, we cannot use this approach. However, Neymann-Pearson have demonstrated that we can substitute the likelihood ratio with the ratio of maxima among the two hypothesis distributions.

To simplify notations, we introduce a N-dimensional space, where each dimension represents a particular detector. In this space, each network vector has N components. So:

$$\begin{aligned} F_+ &= \{F_{1+}, F_{2+}, ..., F_{N+}\} \\ F_\times &= \{F_{1\times}, F_{2\times}, ..., F_{N\times}\} \\ \sigma &= \{\sigma_1, \sigma_2, ..., \sigma_N\} \\ \zeta &= \{\zeta_1, \zeta_2, ..., \zeta_N\} \\ A &= \{A_1, A_2, ..., A_N\} \end{aligned} \tag{3.13}$$

We introduce also the normalized vectors:

$$\begin{aligned} f_+ &= \left\{\tfrac{F_{1+}}{\sigma_1}, ..., \tfrac{F_{N+}}{\sigma_N}\right\} \\ f_\times &= \left\{\tfrac{F_{1\times}}{\sigma_1}, ..., \tfrac{F_{N\times}}{\sigma_N}\right\} \\ A_\sigma &= \left\{\tfrac{A_1}{\sigma_1}, ..., \tfrac{A_N}{\sigma_N}\right\} \end{aligned} \tag{3.14}$$

and the so called network antenna patterns:

$$g_r = \sum_{k=1}^{N} \frac{A_k \cdot \tilde{A}_k}{\sigma_k^2} \tag{3.15}$$

$$g_c = \sum_{k=1}^{N} \frac{A_k^2}{\sigma_k^2} \tag{3.16}$$

We focus on the transformation $g_c \rightarrow g_c'$ which makes the imaginary part of g_c' null. This transformation is useful because it simplifies some equations (some terms become null), as we see in the following considerations. If $g_c = |g_c|e^{2i\gamma}$ the transformation is: $A_k' = A_k e^{-i\gamma}$ and, conseguently,

$$\begin{aligned} F_{k+}' &= F_{k+}Cos(\gamma) + F_{k\times}Sin(\gamma) \\ F_{k\times}' &= -F_{k+}Sin(\gamma) + F_{k\times}Cos(\gamma) \end{aligned} \tag{3.17}$$

The transformation is analogous for the normalized antenna pattern: $A_{\sigma k}' = A_{\sigma k}e^{-i\gamma}$

$$\begin{aligned} f_{k+}' &= f_{k+}Cos(\gamma) + f_{k\times}Sin(\gamma) \\ f_{k\times}' &= -f_{k+}Sin(\gamma) + f_{k\times}Cos(\gamma) \end{aligned} \tag{3.18}$$

A null imaginary part of g_c' is equivalent to the orthogonality of the vectors f_+' and $f_\times'$, as it can be easily shown:

$$\begin{aligned} A_\sigma'^2 &= A_\sigma' \cdot A_\sigma' = \\ &= (f_+' + if_\times') \cdot (f_+' + if_\times') = \\ &= f_+'^2 - f_\times'^2 + 2i(f_+' \cdot f_\times') \end{aligned} \tag{3.19}$$

We call this frame **Dominant Polarization Frame** (**DPF**).

From the previous equation we can define the transformation from a general frame $(f_+, f_\times)$ to the DPF frame $(f'_+, f'_\times)$, i.e. we express the angle γ as a function of $(f_+, f_\times)$.

$$f_+^2 - f_\times^2 + 2i(f_+ \cdot f_\times) = A_\sigma^2 = (A'_\sigma e^{i\gamma})^2 = |A'^2_\sigma| e^{2i\gamma} = |A'^2_\sigma|[Cos(2\gamma) + iSin(2\gamma)] \tag{3.20}$$

From this we have[2]:

$$\begin{array}{c} Cos(2\gamma)|A_\sigma^2| = (f_+^2 - f_\times^2) \\ Sin(2\gamma)|A_\sigma^2| = 2(f_+ \cdot f_\times) \end{array} \tag{3.21}$$

These are the transformation from the generic frame $(f_+, f_\times)$ to the DPF frame. We should calculate from these equalities the expression of $Cos(\gamma)$ and $Sin(\gamma)$, but we are not interested in it, as we will expose now.

We can easily see that these transformations assure the orthogonality of f_+ and $f_\times$ vectors. Infact:

$$\begin{aligned} f'_+ \cdot f'_\times &= (f_+ Cos(\gamma) + f_\times Sin(\gamma)) \cdot (-f_+ Sin(\gamma) + f_\times Cos(\gamma)) = \\ &= -(f_+^2 - f_\times^2) Cos(\gamma) Sin(\gamma) + (f_+ \cdot f_\times)(Cos^2(\gamma) - Sin^2(\gamma)) = \\ &= -(f_+^2 - f_\times^2) Sin(2\gamma)/2 + (f_+ \cdot f_\times) Cos(2\gamma) = 0 \end{aligned} \tag{3.22}$$

We can see that the antenna patterns defined in the DPF are related to those defined in the generic frame (Eq: 3.15 and 3.16).

$$\begin{aligned} |f'_+|^2 &= |f_+ Cos(\gamma) + f_\times Sin(\gamma)|^2 = \\ &= f_+^2 Cos^2(\gamma) + f_\times^2 Sin^2(\gamma) + (f_+ \cdot f_\times) Cos(\gamma) Sin(\gamma) = \\ &= f_+^2 \frac{1 + Cos(2\gamma)}{2} + f_\times^2 \frac{1 - Cos(2\gamma)}{2} + (f_+ \cdot f_\times) \frac{Sin(2\gamma)}{2} = \\ &= \frac{1}{2}((f_+^2 + f_\times^2) + |A_\sigma^2|) = \frac{1}{2}(|A_\sigma|^2 + |A_\sigma^2|) = \frac{1}{2}(g_r + |g_c|) \end{aligned} \tag{3.23}$$

and for $|f'_\times|^2$

$$|f'_\times|^2 = \frac{1}{2}(|A_\sigma|^2 - |A_\sigma^2|) = \frac{1}{2}(g_r - |g_c|) \tag{3.24}$$

and from these relations we obtain an important property of the f'_+, $f'_\times$ vectors

$$|f'_+|^2 \geq |f'_\times|^2 \quad \rightarrow \quad 0 \leq \frac{|f'_\times|^2}{|f'_+|^2} \leq 1 \tag{3.25}$$

and we can make the following considerations.

[2]Obviously, from relation $A'_k = A_k e^{-i\gamma}$ it follows that $|A'_\sigma| = |A_\sigma|$

If $|f'_\times| = 0$ the detectors see the same plus and cross components, as they would have parallel arms. This follows from the inverse of Eq. 3.18, where we obtain:

$$\begin{array}{c} f_{k+} = f'_{k+} Cos(\gamma) \\ f_{k\times} = f'_{k+} Sin(\gamma) \end{array} \tag{3.26}$$

i.e. the ratios between plus and cross components are the same for all detectors.

Moreover, if $|f'_+| = |f'_\times|$, this is valid for a generic frame. Infact from Eq. 3.24 it follows that $|A^2_\sigma| = 0$ and this conditions is satisfied if the real and imaginary part of Eq. 3.19 are null:

$$\begin{array}{c} f^2_+ = f^2_\times \\ f_+ \cdot f_\times = 0 \end{array} \tag{3.27}$$

i.e. the antenna patterns in the general frames are orthogonal and with the same module.

In the DPF, the likelihood assumes a simple expression.[3]
If we define the normalized vector: $X = \left\{ \frac{x_1}{\sigma_1}, ..., \frac{x_n}{\sigma_N} \right\}$ and $\xi_\sigma = \left\{ \frac{\xi_1}{\sigma_1}, ..., \frac{\xi_n}{\sigma_N} \right\}$, the likelihood becomes:

$$\begin{aligned} L &= \left(X \cdot \xi_\sigma - \frac{1}{2} \xi_\sigma \cdot \xi_\sigma \right) = \\ &= \left[X \cdot (f'_+ h_+ + f'_\times h_\times) - \frac{1}{2} (f'_+ h_+ + f'_\times h_\times) \cdot (f'_+ h_+ + f'_\times h_\times) \right] = \\ &= \left[X \cdot f'_+ h_+ + X \cdot f'_\times h_\times - \frac{1}{2} (f'^2_+ h^2_+ + f'^2_\times h^2_\times) \right] \end{aligned} \tag{3.28}$$

where the last relation holds only in the DPF. The maximum likelihood is calculated from the first derivatives $\frac{\delta L}{\delta h_+} = 0$ and $\frac{\delta L}{\delta h_\times} = 0$:

$$\begin{array}{l} \frac{\delta L}{\delta h_+} = X \cdot f'_+ - |f'_+|^2 h_+ = 0 \\ \frac{\delta L}{\delta h_\times} = X \cdot f'_\times - |f'_\times|^2 h_\times = 0 \end{array} \tag{3.29}$$

which give the solutions:

$$\begin{array}{c} h_+ = \frac{X \cdot f'_+}{|f'_+|^2} \\ h_\times = \frac{X \cdot f'_\times}{|f'_\times|^2} \end{array} \tag{3.30}$$

[3]For the sake of simplicity we do not explicitily write $\sum_i$.

The maximum likelihood is obtained by introducing such values in the Eq. 3.28:

$$\begin{aligned} L_{max} &= \left[X \cdot f'_+ h_+ + X \cdot f'_\times h_\times - \frac{1}{2}(f'^2_+ h^2_+ + f'^2_\times h^2_\times) \right] = \\ &= \left[X \cdot f'_+ \frac{X \cdot f'_+}{|f_+|^2} + X \cdot f'_\times \frac{X \cdot f'_\times}{|f'_\times|^2} - \frac{1}{2} \left(f'^2_+ \frac{(X \cdot f'_+)^2}{|f_+|^4} + f'^2_\times \frac{(X \cdot f'_\times)^2}{|f'_\times|^4} \right) \right] = \\ &= \left(\frac{(X \cdot f'_+)^2}{|f'_+|^2} + \frac{(X \cdot f'_\times)^2}{|f'_\times|^2} \right) \end{aligned} \tag{3.31}$$

So, the detector response of the solution is:

$$\begin{aligned} \xi_\sigma &= h_+ f'_+ + h_\times f'_\times \\ &= \frac{X \cdot f'_+}{|f'_+|^2} f'_+ + \frac{X \cdot f'_\times}{|f'_\times|^2} f'_\times \\ &= e'_+ \frac{X \cdot f'_+}{|f'_+|} + e'_\times \frac{X \cdot f'_\times}{|f'_\times|} \end{aligned} \tag{3.32}$$

where e'_+ and $e'_\times$ are the unitary vector $e'_+ = \frac{f'_+}{|f'_+|}$ and $e'_\times = \frac{f'_\times}{|f'_\times|}$. The detector response is the projection of the x vector on the plane defined by f'_+ and $f'_\times$ and the Likelihood Ratio it equivalent the module of this projection (see Fig. 3.1);

$$L_{max} = \frac{(X \cdot \xi_\sigma)^2}{|\xi_\sigma|^2} \tag{3.33}$$

We can see that the maximum likelihood is related to the SNR values of the single detectors. Considering the detector k, the SNR is calculated from detector response ξ_k and noise variance σ_k: $\mathrm{SNR}^2_k = \frac{\xi^2_k}{\sigma^2_k}$. We can write the square sum of detector SNR with the vector notation:

$$\sum_k \mathrm{SNR}^2_k = \sum_k \frac{\xi^2_k}{\sigma^2_k} = \sum_k \xi^2_{\sigma k} = \xi_\sigma \cdot \xi_\sigma \tag{3.34}$$

The two vectors ξ_σ and X are related: $X = \xi_\sigma + n$, where n is the noise, and we can write:

$$X \cdot \xi_\sigma = \xi_\sigma \cdot \xi_\sigma + n \cdot \xi_\sigma = \xi_\sigma \cdot \xi_\sigma \tag{3.35}$$

where we have applied the orthogonality between ξ and n. Using this definition we can find the relation between detector SNR and likelihood:

$$L = \frac{(X \cdot \xi_\sigma)^2}{|\xi_\sigma|^2} = \frac{(\xi_\sigma \cdot \xi_\sigma)^2}{|\xi_\sigma|^2} = \xi_\sigma \cdot \xi_\sigma = \sum_k \mathrm{SNR}^2_k \tag{3.36}$$

which is the expression of maximum likelihood as in Eq. 3.33. The maximum likelihood is the sum of detector square SNRs.

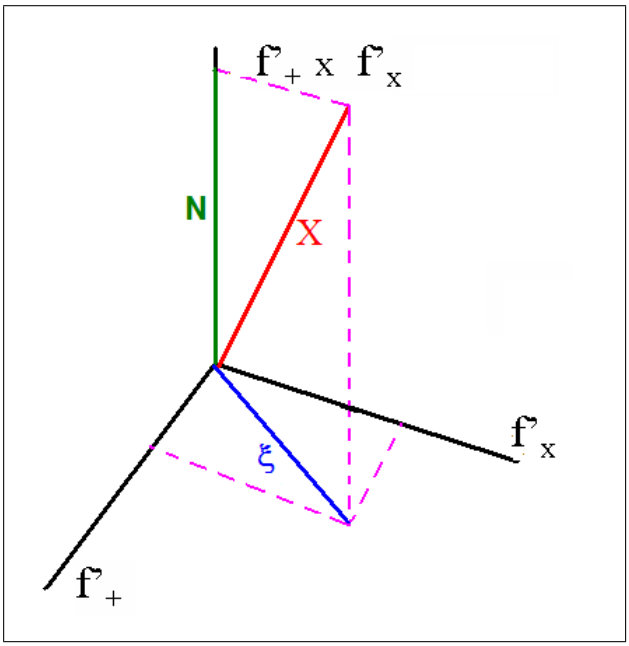

Figure 3.1: *Dominant Polarization Frame*

3.1.3 Network Antenna Patterns

We just esplained how Waveburst combine the detector antenna pattern to construct the network antenna pattern. In the Fig. 2.5 we report the $sqrt|f'_+|^2 + |f'_\times|^2$ quantity of the single detectors L1, H1, V1, which are the one we are more interested, because they are considered in this thesis. Obviously, we do not report the H2 because its antenna patterns are the same as H1.

In the Fig 3.2 we report the module of the plus component considering the three couples of detectors. We can see that the more sensitive zones (red colour) increases with respect to the single detectors. This smartly shows the advantage to use more than one detector.

Moreover, we report in Fig. 3.3 the same quantity for the network including all the available detectors[4] and the ratio $|f'_\times|/|f'_+|$. We see that the plus component is always greater than the cross component, and that in most of ths sky the cross component is small with respect to the plus component (blue zones in the bottom of the figure). This is equivalent to the situation of $f'_\times$ is almost null.

[4]We remind that we do not consider GEO600 because its sensitivity is not as the same level of the other detectors

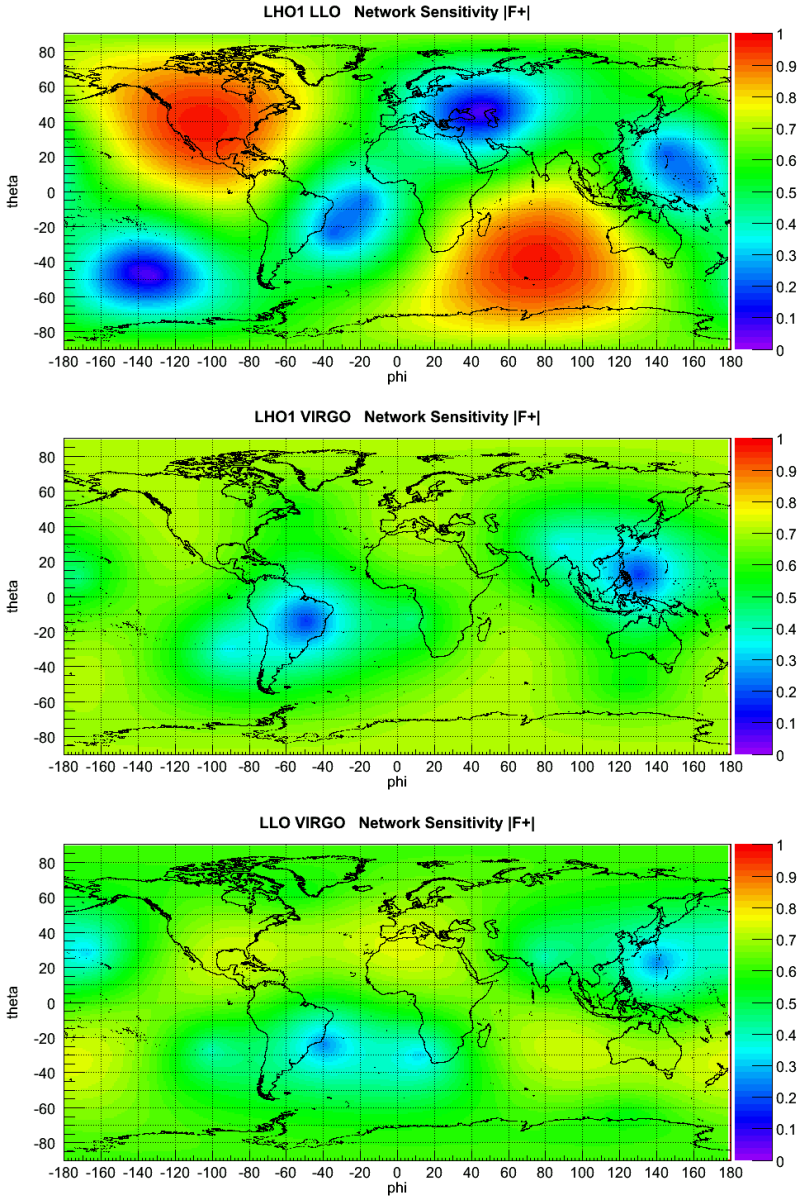

Figure 3.2: *Variation of plus component of antenna patterns following earth coordinate for a network of two detectors. On x-axis is reported the longitude, and on y-axis the latitude. Coloured axis reports valued of antenna patterns. From top to bottom the networks are: H1L1, H1V1, L1V1.*

3.1.4 Regulators

We have defined the maximum likelihood starting from the definition of useful quantities: ξ_σ, f_+, $f_\times$, $L = (X \cdot \xi_\sigma)^2/|\xi_\sigma|^2$. We have assumed that the noise is

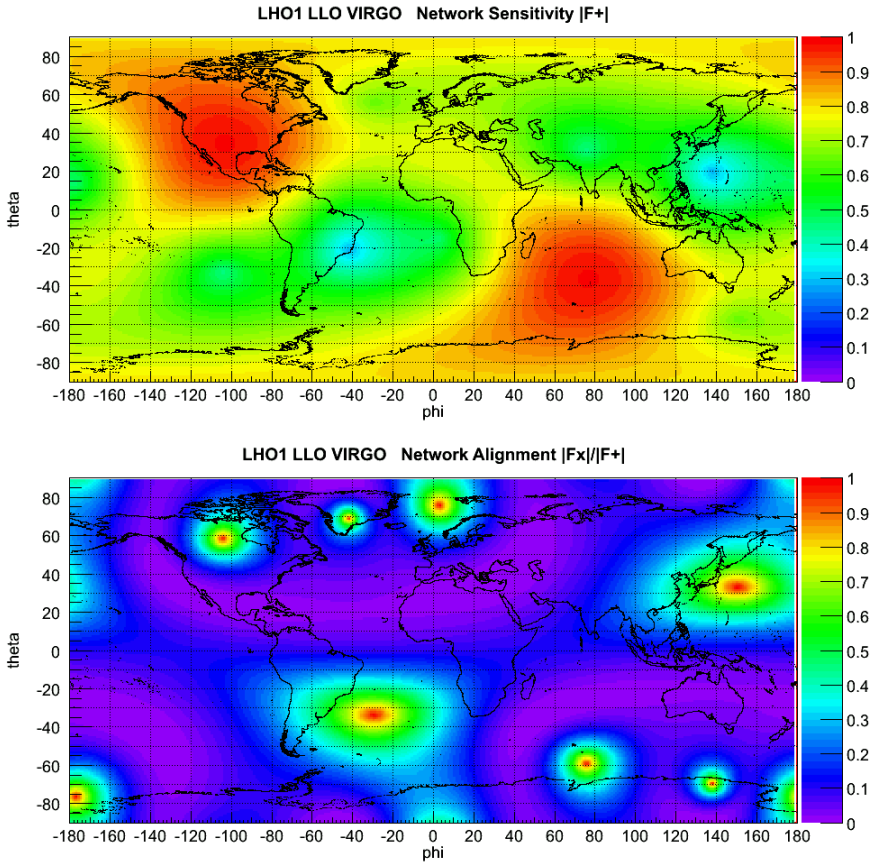

Figure 3.3: *Variation of L1H1V1 network antenna patterns in the DPF following earth coordinate:* $|f'_+|$ *on the top and* $|f'_\times|/|f'_+|$ *on the bottom. On x-axis is reported the longitude, and on y-axis the latitude. Coloured axis reports valued of antenna patterns.*

on the sub-space orthogonal to the plane $(f'_+,\ f'_\times)$, but a part of the noise could be also on this plane. Referring to fig. 3.4, the blue vector is the projection of x vector on the plane, which we suppose disturbed by noise, and the brown vector is the corrected one not affected by the residual noise (the dashed green line).

To address this problem we introduce the **Regulators**: i.e. some assumptions on the network properties, with the purpose to exclude unphysical solutions due to the variation of the likelihood functional.

The idea is to make statystical assumptions about the position of the detector response on the $(f'_+, f'_\times)$ plane and it assumes an uniform distribution varying the polarization angle. A gravitational wave, in fact, could assume any possible position on this plane. The detector response, however, is affected from the antenna patterns, so it would be distributed preferrable on the component with the greater

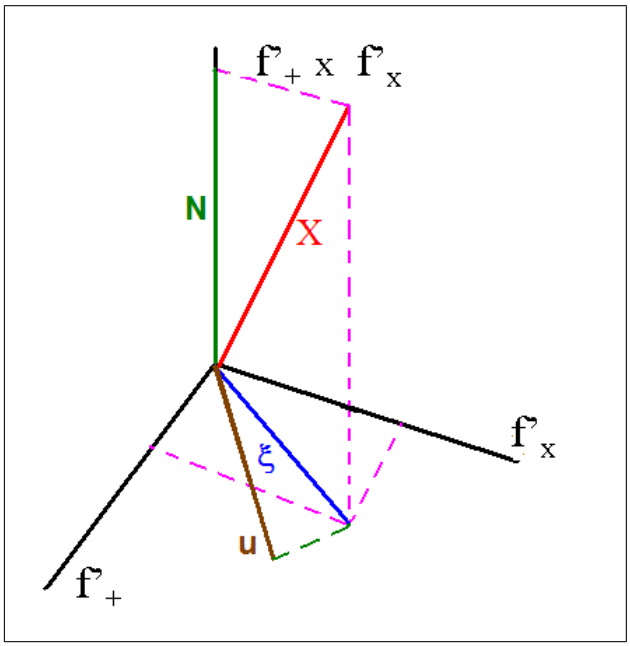

Figure 3.4: *Detector response affected by noise (blue) and solution not affected by noise (brown) in the Dominant Polarization Frame*

module of the antenna pattern (In this case the plus component, according to how we have defined the Dominant Polarization Frame).

These considerations are visually esplained in Fig. 3.5. If we consider generic wave with an circular polarization, its distribution on the plane would be uniform for the angle respect to plus component, and constant in amplitude. The detector response, however, depends on the antenna pattern, and its amplitude depend on the angle, and the angles would not be more uniform.[5]

So, the space of possible solutions is limited, following the assumptions on the network. The goal is to improve the detection for the most part of the sources expecting to loose a small fraction of the real GW signals [64].

For the definition of the regulator is it useful to introduce a new vector on the $(f'_+, f'_\times)$ that it is suggested from Eq. 3.33. The idea is to define u so to substitute it to the ξ_σ vector on the Maximum Likelihood:

$$L_{max} = \frac{(X \cdot u)^2}{|u|^2} \tag{3.37}$$

[5]Consider, for instance, the position with angle $\pi/4$. This satysfies $h_+ = h_\times$, but the corresponding position of $\xi = h_+ f'_+ + h_\times f'_\times$ has an angle with the f'_+ axis depending on the ratio $|f'_\times|/|f'_+|$.

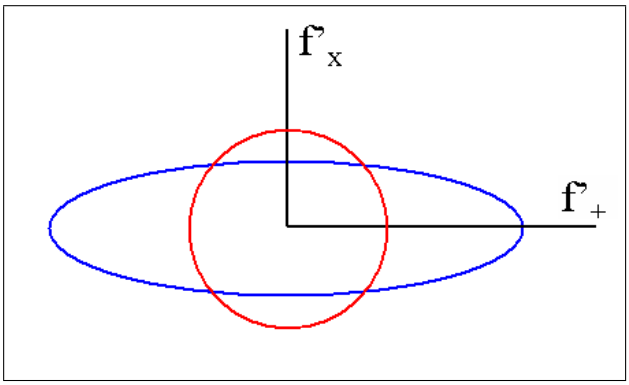

Figure 3.5: *Visualization of a circular polarized signal (red) and of the related detector response (blue) in the Dominant Polarization Frame. Plus component is greater because of the definition of DPF.*

The u vector is an useful tool because the normalization for $|u|$ allows us to disregard u module and to be interested only in its direction. With this vector it is also possible to calculate likelihood also in any generic frame, without calculating vectors in the DPF. We see in the Appendix B how it is possible.

Regulators are defined by changing the definition of u vector. The possible choices in coherent Waveburst are:

- **weak**: $u_{weak} = (X \cdot e'_+)e'_+ + (X \cdot e'_\times)e'_\times$
- **soft**: $u_{soft} = (X \cdot f'_+)f'_+ + (X \cdot f'_\times)f'_\times$
- **hard**: $u_{hard} = f'_+$
- **mild**: $u_{mild} = f'_+ \pm f'_\times$

The choice of the $\pm$ sign for mild regulator depends on where the projection of X lies in the $(f'_+, f'_\times)$ plane, if it is on the first/third or second/fourth quadrant. The weak regulator corresponds to the likelihood approach we have previously introduced, the hard regulator imposes the signal has only the plus component, instead the others are intermediate situations between weak and hard depending on the ratio between $|f'_\times|/|f'_+|$.

Considering the soft definition, if $|f'_\times|/|f'_+| = 0$, this regulator becomes obviously the hard one. However, if $|f'_\times|/|f'_+| = 1$ we can write:

$$\begin{aligned} u_{soft} &= (X \cdot f'_+)f'_+ + (X \cdot f'_\times)f'_\times = \\ &= |f'_+|^2(X \cdot e'_+)e'_+ + |f'_\times|^2(X \cdot e'_\times)e'_\times = \\ &= |f'_+|^2 \; u_{weak} \end{aligned} \tag{3.38}$$

Also the mild regulator becomes the hard one when the $|f'_\times|/|f'_+| = 0$. Moreover, when $|f_+| \gg |f_\times|$ it is approaches the weak regulator. In fact, the network response ξ depends on the antenna patterns, so, if $|f_+| \gg |f_\times|$ also the network response $\xi \approx h_+ f_+$, and mild regulators approx weak one. Only when $|f_+|$ and $|f_\times|$ are comparable, weak and mild regulator can differ substantially.

Unfortunately, in general, there is no optimal rule to select the right regulator. A chosen regulator is better for certain situations but can be worst for others. For instance, the *hard* regulator imposes that the $|f'_\times| = 0$, which is equivalent to assume that the GW has only the h_+ component, or that the source direction is such that the network is not sensible to the $h_\times$ polarization. This is certainly true for some directions in the sky, but there are also certain directions for which $|f'_\times| \approx |f'_+|$ and the hard regulator is therefore not efficient.

However, the role of regulators is to obtain a good efficiency and a low detection of glitches (and a good estimation of source direction, of course!). Different regulators assumes several signal characteristic that provide better/worse efficiency and contemporanealy increase/decrease glitches detection. For instance, the *weak* regulator is the one that better fits the signal, because for each pixel it projects data vector on network antenna pattern plane. However, for the same reason, such feature causes an huge number of detected glitches, because of its capability to adapt to the signal characteristic, noise included. When we use a network of two detectors, the algorithm is not able to find the right position on the $(f'_+, f'_\times)$ plane, so it is necessary to use a regulator, more specifically the hard one, that for most of the direction in the sky is a good approximations (see Fig. 3.2).

Hard regulator uses more limited assumptions and glitches detection is therefore lower than weak regulator, but, as we have already realized for some situations, this is not a good feature. Also *mild* regulator effects on glitches detection are more stringent than weak. This is due to the fact the mild regulator imposes that the detector response is affected by the antenna pattern, so it is probable that its direction in the $(f'_+,\ f'_\times)$ plane is affected by the reciprocal module ratio.

3.1.5 Energy disbalance

Regulators use information about the network and obviously they are not enough to find the solution of the likelihood calculation, but they are only a starting point. To find a more realistic solution, we take in account some property that follow from the likelihood approach, introducing a new quantity: the **Energy Disbalance**.

We remind the expression of the likelihood:

$$L = \sum_k \left(\frac{x_k^2}{\sigma_k^2} - \frac{(x_k - \xi_k)^2}{\sigma_k^2} \right) \tag{3.39}$$

and the following definitions:

$$f_{k+} = \frac{F_{k+}}{\sigma_k} \qquad f_{k\times} = \frac{F_{k\times}}{\sigma_k} \qquad X = \frac{x_k}{\sigma_k} \tag{3.40}$$

With this notations the normalixed detector response can be written: $\xi_\sigma = f_+ h_+ + f_\times h_\times$, where the two vectors ξ_σ e X are related to the detectors noise ($X = \xi_\sigma + n$).

We now consider the *null stream*:

$$n_k^2 = (X_k - \xi_{\sigma k})^2 = X_k^2 - \xi_{\sigma k}^2 - 2(X_k \xi_{\sigma k} - \xi_{\sigma k}^2) \tag{3.41}$$

where we call the last term *detector energy disbalance*

$$\Delta_k = X_k \xi_{\sigma k} - \xi_{\sigma k}^2 \tag{3.42}$$

From the likelihood analysis it follows that:

$$L_{max} = \sum X_k \xi_{\sigma k} = \sum \xi_{\sigma k}^2 \tag{3.43}$$

so the total energy disbalance $\sum \Delta_k$ is equal to zero. This follows from the likelihood analysis, in fact, the likelihood calculation impose the minimization of the null stream. The total energy disbalance is the scalar product between the null stream and the detector response:

$$\sum \Delta_k = \sum_k X_k \xi_{\sigma k} - \xi_{\sigma k}^2 = (X - \xi_\sigma) \cdot \xi_\sigma = n \cdot \xi = 0 \tag{3.44}$$

Such results does not guarantee that for each individual term $\Delta_k = 0$, and unconstrained likelihood variation may result in unphysical solutions when $\Delta_k \ll 0$. In this case the energy of reconstructed detector response ξ_σ^2 can be much greater than the energy of the detector data stream X^2.

To face this problem we consider the solution that minimizes the quantity $\sum_k (X_k \xi_{\sigma k} - \xi_{\sigma k}^2)^2$. From Fig. 3.6 we start from the position in the $(f'_+,\ f'_\times)$ defined by the chosen regulator (u) and we find the nearest position which has a minimum of energy disbalance quantitity (u').

We therefore define the *energy disbalance*:[6]

$$q_k = \frac{X_k \xi_{\sigma k} - \xi_{\sigma k}^2}{h} = X_k u_k - u_k^2 h \tag{3.45}$$

which depends on the u vector. We can express this dependence considering the angle ψ defined by the two vectors u and u_0, where u_0 is the direction selected by

[6]We decide to divide the quantity $X_k \xi_k - \xi_{\sigma k}^2$ to h because in this way we do not consider the solution: $h = 0$

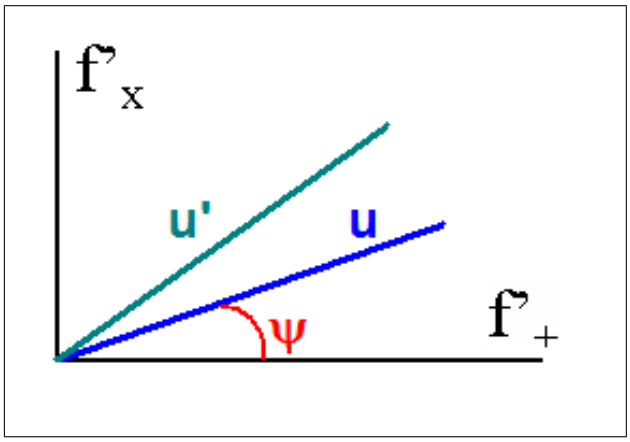

Figure 3.6: *Position of the u vector of a regulator and of the associated u′ vector given from the minimization of the energy disbalance, ψ is the angle between the vector and the f'_+ axis.*

the regulator. Then we introduce the vector $v = \frac{\partial u}{\partial \psi}$ (orthogonal to u) and we can write:

$$\begin{aligned} u &= u_0 Cos(\psi) + v_0 Sin(\psi) \\ v &= -u_0 Sin(\psi) + v_0 Cos(\psi) \end{aligned} \tag{3.46}$$

where v_0 is the orthogonal direction to u_0.

We find the minimum of the expression: $\frac{\partial}{\partial \psi}(q \cdot q) = 0$. We define the vector p such as $\frac{\partial}{\partial \psi}(q \cdot q) = -2q \cdot p$, i.e.:

$$\begin{aligned} q_k &= X_k u_k - u_k^2 h \\ p_k &= (2u_k h - X_k)v_k + u_k^2 (X \cdot v)/|u|^2 \end{aligned} \tag{3.47}$$

Infact

$$\begin{aligned} p_k &= -\frac{\partial q_k}{\partial \psi} = -\frac{\partial}{\partial \psi}(X_k u_k - u_k^2 h) \\ &= -\left[X_k \frac{\partial u_k}{\partial \psi} - 2u_k h \frac{\partial u_k}{\partial \psi} - \frac{u_k^2}{|u|^2}\left(X \cdot \frac{\partial u_k}{\partial \psi} \right) \right] \\ &= -[X_k v_k - 2u_k h v_k - u_k^2 (X \cdot v)/|u|^2] = (2u_k h - X_k)v_k + u_k^2 (X \cdot v)/|u|^2 \end{aligned} \tag{3.48}$$

We normalize the disbalance for the quantity $\frac{X_k}{X_k^2 + \delta}$ to let the detectors in the network be comparable (otherwise a detector with a big disbalance could be present or a small SNR could modify consistently the minimum).

If we assume $\psi \sim 0$, the definition of q and p can be approximated with: (see Appendix C)

$$\begin{array}{c} q \sim q_0 Cos(\psi) - p_0 Sin(\psi) \\ p \sim q_0 Sin(\psi) + p_0 Cos(\psi) \end{array} \tag{3.49}$$

where $q_0 = q(u_0, v_0)$ and $p_0 = p(u_0, v_0)$. Using these approximations, the condition $q \cdot p = 0$ is satisfied by:

$$\begin{array}{c} Cos(2\psi_{min}) \sim p_0^2 - q_0^2 \\ Sin(2\psi_{min}) \sim 2(q_0 \cdot p_0) \end{array} \tag{3.50}$$

The angle ψ_{min} is related to a minimum of $q \cdot q$, as it can be easily shown by calculating the second derivative:

$$\begin{aligned} \frac{\partial^2}{\partial \psi^2}(q \cdot q) &= -2\frac{\partial}{\partial \psi}(q \cdot p) = \\ &= -2\left(\frac{\partial q}{\partial \psi} \cdot p + \frac{\partial p}{\partial \psi} \cdot q\right) = \\ &\sim -2\left(-p \cdot p + q \cdot q\right) \end{aligned} \tag{3.51}$$

where we have used the approximated definition of $p \approx q_0 Sin(\psi) + p_0 Cos(\psi)$, from which we find $\frac{\partial p}{\partial \psi} = \frac{\partial}{\partial \psi}(q_0 Sin(\psi) + p_0 Cos(\psi)) = q_0 Cos(\psi) - p_0 Sin(\psi) = q$. The second derivative becomes:

$$\begin{aligned} p \cdot p - q \cdot q &= (q_0^2 Sin^2(\psi) + p_0^2 Cos^2(\psi) + 2q_0 \cdot p_0 Cos(\psi) Sin(\psi)) + \\ &- (q_o^2 Cos^2(\psi) + p_0^2 Sin^2(\psi) - 2q_0 \cdot p_0 Cos(\psi) Sin(\psi)) = \\ &= 4q_0 \cdot p_0 Cos(\psi) Sin(\psi) + (p_0^2 - q_0^2)(Cos^2(\psi) - Sin^2(\psi)) \end{aligned} \tag{3.52}$$

Using Eq. 3.50 to calculate this expression for $\psi = \psi_{min}$

$$\begin{aligned} (p \cdot p - q \cdot q)|_{\psi_{min}} &= (2q_0 \cdot p_0 Sin(2\psi) + (p_0^2 - q_0^2) Cos(2\psi))|_{\psi_{min}} = \\ &= 4(q_0 \cdot p_0)^2 + (p_0^2 - q_0^2)^2 = \\ &= (p_0^2 + q_0^2)^2 > 0 \end{aligned} \tag{3.53}$$

Secund derivative is always positive, so ψ_{min} refers to a minimum.

We can therefore calculate the maximum likelihood, defined as the projection of X on u:

$$L = \frac{(X \cdot u(\psi_{min}))^2}{|u(\psi_{min})|^2} \tag{3.54}$$

3.1.6 Assumptions on the signal

Up to now, we did not make any assumption on the signal characteristics. Waveburst includes the possibility to use a priori assumptions on the signal when calculating the maximum likelihood. In particular, these assumptions regard the signal polarization, which could be random (no assumptions), elliptical, circular, or linear. Following this order, the algorithm becomes more rescrictive, so the set of targeted signal decreases. As for regulators, if we impose more rescrictive conditions, glitches decreases and efficiency detection depend more strictly on the waveform properties. So, if we use the linear assumption we expect that detection for signals with random polarization would be very low.

We use the word **constrain** to refer to an assumption on the signal polarization. Even if we do not make assumptions, we will say that we use the *unmodelled* constrain. For the other assumptions, we use the polarization type (*elliptical*, *circular*, *linear*).

To identify the different constrains, we introduce labels, which are listed in the table 3.2

Polarization	**Label**
unmodelled	B
linear	S
circular	G
elliptical	I

Table 3.2: *Constrains used in Coherent Waveburst and their relative label.*

To apply a constrain, the algorithm considers the signal composed by two othogonal components, that we call X_{00} and X_{90}. Then the likelihood function is calculated for both the components along the sky locations. The maximum likelihood over the sky is calculated combining the L_{00} and L_{90} according to the polarisation type considered. For the unmodelled constrain, only the X_{00} and L_{00} are considered.

3.2 Production algorithms

In this section we describe with more details the production stage of the CWB pipeline. CWB starts from the raw data and produces a list of triggers which are saved as output in a ROOT file. ROOT [65] is a library collection based on the C++ language developed by the CERN for the High Energy Analysis. The ROOT structure is very comfortable also for the GW Data Analysis, thanks to

their graphycal and mathematical libraries. Moreover, ROOT includes a command interpreter which allows to run macros in the interactive way, i.e. without compyling them. This characteristic makes the instrument adjustable also to faster and not complicated analysis and allows to make a large variety of small tests.

The likelihood method introduced in the previous paragraph requires high computational load and use of memory, so a great effort has been used to introduce useful approximations or matematical tricks to makes the analysis faster and with low computational load. For instance, it is possible to run multiple analyses (**jobs**) on different time periods (**segments**) exploiting the use of more machines. Time segments have a limited length, usually between 300 and 600 s. The lower limit is established taking in account that too small periods can introduce dangerous approximations (especially on the Wavelet Transform, see the next section), where as the upper limit is chosen to allows reduced computational load.

The algorithm allows the possibility to receive in input a list of allowed times which excludes more noisy periods included in the segments which it is better to avoid for the data analysis. For more details, see the Sec. 5.1.1.

Each job produces a respective ROOT file, all the events included in ROOT files referring to the same analysis can be collected in a unique ROOT file (*merging procedure*).

The production stage of the pipeline can be summarized in the following list:

- Data conditioning
- Wavelet transformation
- Multi-Resolution analysis
- Time-shift analysis
- Calculation of the likelihood over the sky
- Reconstruction of waveform parameters

3.2.1 Data conditioning

The pipeline is designed to run on calibrated data (called $h(t)$). Running the pipeline on raw data would affects negatively the detection efficiency and reconstruction parameters. So it is required to apply the analysis on calibrated data, previous version of Waveburst include the possibility to use not calibrated data, but this procedure is not used [66].

This part includes the following steps:

- **Linear Prediction Error (LPE) filter**. This filter is used to remove "predictable" components from a time series, for instance lines of stationary noise. The LPE filter is used in wavelet domain individually for each wavelet layer. The procedure has been just developed in the incoherent version [67, 58].

- **Whitening**. The standard incoherent Waveburst procedure is applied to produced whitened wavelet series [67, 68]. The whitening is performed for the different wavelet layers and should be used for the likelihood calculation. Applying inverse wavelet transform to normalized wavelet amplitudes r_{ij} the pipeline produce whitened time series.

3.2.2 Wavelet transform

The main idea of the Waveburst algorithm is to use the Wavelet Transform [69], an useful tool that allows to describe the signal both in time and frequency at the same time. The advantages to introduce Wavelet Transform are:

- on the possibility to use several time-frequency resolutions to find an optimal characterization of the signal

- on the use of an orthogonal representation, i.e. complete and not redundant

- on low computational usage.

The Discrete Wavelet Transform (DWT) is a linear combination of a chosen function (called *mother wavelet*) which is rescaled and traslated according to a sampling procedure applyed on the Continue Wavelet Transform. The sampling grid is defined as:

$$\begin{cases} s = 2^{-j} & j \in \mathbb{Z} \\ \tau = n2^{-j} & n \in \mathbb{Z}, j \in \mathbb{Z} \end{cases} \tag{3.55}$$

The DWT of the function $x(t)$ is defined as:

$$DWT(x) = \sum_{s,\tau} d_{s,\tau} \psi_{s,\tau}(t) \tag{3.56}$$

where we have defined:

$$\psi_{s,\tau}(t) = \frac{1}{\sqrt{s}} \psi \left(\frac{t-\tau}{s} \right). \tag{3.57}$$

from the wavelet mother $\psi(t)$. This function should satisfyes admisibility condition[7] so that the inverse wavelet transform is equal to the original function. The

[7] $C_\psi = \int_0^{+\infty} \frac{|\Psi(\omega)|^2}{|\omega|} d\omega$ should be a finite value, where $\Psi(\omega)$ is the Fourier Transform of $\psi(t)$

$d_{j,n}$ coefficient are defined from:

$$d_{j,n} = \int_{-\infty}^{+\infty} x(t)\psi^*_{j,n}(t)dt, \tag{3.58}$$

The couple (s, τ) defines a time-frequency pixel, i.e. a time interval and a frequency band centered in definite time and frequency. The square of coefficient $d(s, \tau)$ is the energy of the signal related to the time-frequency pixel which is associated to the couple (s, τ).

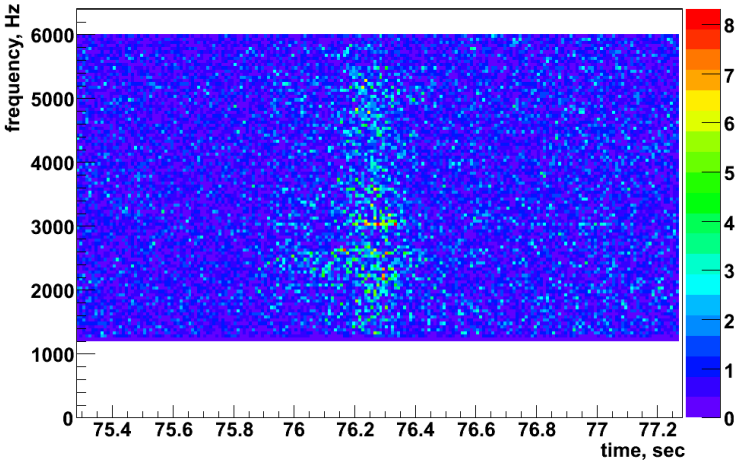

Figure 3.7: *Example of a Wavelet Tranform on the data stream selecting the frequency band [1280, 6000] Hz. The detector considered is V1.*

The time-frequency resolutions (i.e. the largeness of the time interval and frequency band) are related to the couple (s, τ). It is not possible to choose an arbitrary resolution, this quantity depends on the decomposition level used and the central frequency. However, it is possible to use linear combination of wavelet functions to have the same time-frequency resolution at all frequencies: this procedure is called Wavelet Packet. In this way the signal is characterized at different decomposition levels and each level has a proper time-frequency resolution. Each decomposition level is an indipendent and complete characterization of the original signal. Passing from one level to the next, time resolution doubles and frequency resolution halves.

3.2.3 Multi-Resolution Analysis

Waveburst uses different decomposition levels for the same data stream, such as to have different characterization of the signal and find the optimal one. The

algorithm considers which is the optimal decomposition level according to the characteristic of the signal and calculates the likelihood on this optimal level.

After the application of the Wavelet Transform, the algorithm selects the most energetic pixels (*core*) and their neighbours[8] (*halo*). Core pixels are chosen if the corresponding energy is more than a threshold which depends on the noise level. This threshold in fact is adapted according to the total energy of the wavelet decomposition. This is quite equivalent to choose a percentage of the most energetic pixel for each level, and this percentage is quite constant among the levels.

Core and halo pixels are combined in a cluster, which represents an event reconstructed by a detector at the related time-frequency resolution. The clusters belonging to different decomposition levels are combined in super-clusters and on these super-cluster the algorithm applyes the Likelihood calculation.

The likelihood method can be applied to the different levels of the multi-resolution analysis: pixel, cluster, multi-cluster. So, we can perform several checks at each level verify the coherence of the single components. The general Eq. 3.12 can be written in the way:

$$L_c = \sum_{ij} L(i,j,\theta,\phi) \tag{3.59}$$

where i,j refers to the time-frequency position and θ,ϕ to the source coordinate. The single term can be written:

$$L(i,j,\theta,\phi) = \sum_{k=1}^{N} \frac{1}{\sigma_k^2(i,j)} \left(x_k(i,j,\tau_k)\xi_k(i,j,\theta,\phi) - \frac{1}{2}\xi_k^2(i,j,\theta,\phi) \right) \tag{3.60}$$

The likelihood approach is applicable to the functional above, so we can find the maximum likelihood statistic for a given time location (i, j), maximiming likelihood functional over the source coordinates θ and ϕ ($L_m(i,j)$). The set of these maximum likelihood values for each time location as a function of time and frequency gives the likelihood time-frequency map (LTF).

From this pixels we define coherent clusters, which are made of that pixels which have a maximum likelihood $L_m(i,j)$ greater than a chosen threshold, and it is composed of pixels belonging to all detectors involved in the network.

Final clusters are used to reconstruct the gravitational wave.

3.2.4 Time-shift analysis

In order to estimate background, the pipeline apply relative time shifts to increase statystic. Each time shift is considered as a separate analysis which produce

[8]The neighbours of a pixel are that ones that have in common a side or a vertex with this pixel.

its indipendent trigger, but the pipeline is optimized to implement at the same time more analyses with different time shifts without increasing too much the computational load.

This goal is obtained shifting only one detector: the *reference detector*. Usually a time shift T_k applied to data stream is characterized by the time step t_s and the lag number k: $T_k = kt_s$. The maximum time shift Kt_s depends by the minimum duration of a data segment processed (K is the total number of time lags applied). For CWB the time shift step should be selected to be proportional to the minimum time resolution Δt_j among the TF maps of the considered decomposition level: $t_s = m\Delta t_{j\ \min}$, m is an integer. For each lag number k, reference detector data stream is shifted with respect to the others by the time T_k. In case when the index $i + mk$ runs outside of the data segment selected for the analysis, the pipeline perform cycle time shifts starting from the beginning of the segment. The time shifts are applied circularly, this means that, applyed a time shift equal to T_s, the last T_s seconds on the buffer are placed at the beginning, so no data stream is lost.

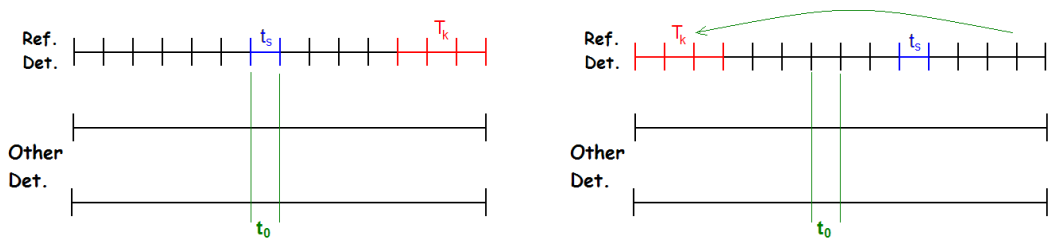

Figure 3.8: *Example before (left) and after (right) applying Time Shift on the reference detector (black line on the top). Blue interval is the time step t_s, red interval is the time shift T_k. In this case $k = 3$.*

This multiple time shifts are performed only for the reference detector. It is possible to apply a **constant** time shift δT_k for each detector, but this cannot be modified during the CWB run.

3.2.5 Calculation of the likelihood over the sky

The likelihood depend on the antenna pattern, so its value varyes according to the position of the sky where we suppose there is the search. For each cluster, the likelihood is calculated for all the sky positions, to locate what is the most probable location for that related event. Obviously, different directions implies different time delay between the detectors, so it is necessary a procedure to apply short time delay to data stream.

It is preferable to calculate time delays in wavelet domain, to simplify calculation of network likelihood matrix. However, if the detector data stream is delayed, the distribution of wavelet amplitude is different from the original one. If we consider a time shift τ applied to the data stream, the TF pixel (i, j) of a certain wavelet layer and detector k assumes the value $x_k(i, j, \tau)$ according to:

$$x_k(i, j, \tau) = \sum_{lp} D_{lp}(\tau, j) x_k(i + l, j + p) \tag{3.61}$$

where $D_{lp}(\tau)$ are the *delay time filters*, which are constructed individually for each wavelet layer. The delay filter construction procedure is the following:

- create blank wavelet series with only one coefficient at the TF location (i, j) set to unity
- apply the inverse wavelet transformation, which will reconstruct a wavelet function $\Phi_j(t)$ in time domain
- shift $\Phi_j(t)$ by time delay τ
- do wavelet decomposition of $\Phi_j(t + \tau)$
- the resulting wavelet amplitudes at TF locations (i+l, j+p) give the delay filter coefficients $D_{lp}(\tau, j)$ for the wavelet layer j.

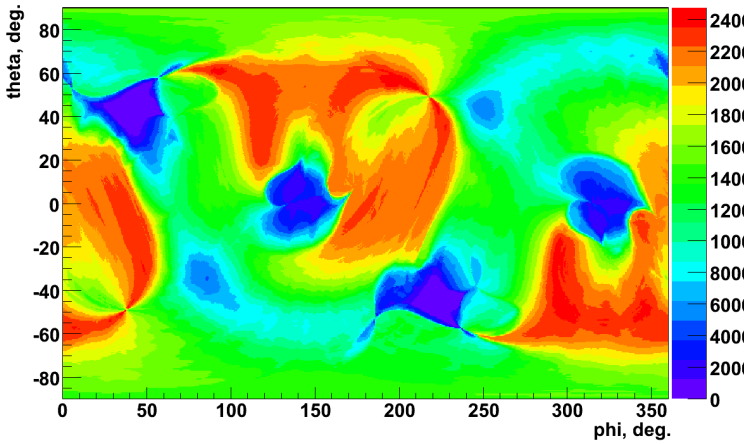

Figure 3.9: *Example of the likelihood behaviour as a function of sky coordinates for an event reconstructed from a background analysis.*

To reduce computational load, delay time filters are applied to the not reference detectors. This, in fact, allows to use the same delays coefficient for all the lags on the reference detector. More details on App. A.1

3.2.6 Reconstruction of waveform parameters

Waveburst uses this likelihood approach to give a complete reconstruction of the waveform. When we use the word *reconstruction* we mean that the algorithm is able to estimate all the information related to the wave, such as: time, frequency, hrss, SNR, and source coordinates.

The reconstruction of waveform parameters can be divided in two principal aspects: the waveform and the source coordinates.

The **waveform reconstruction** is performed using the properties of the Wavelet Transform. In principle, the desired goal should be to reconstruct the h(t) of the GW waveform, but this is not an easy problem, because it requires getting inverse solution of the likelihood equation. To avoid this problem, the algorithm reconstruct the response of each detector, which is easier to calculate. So, we obtain N reconstruction, each referred to the related network.

From any of the used decomposition levels we can always apply the Inverse Wavelet Transform to obtain the original signal in the time domain. Once made a cluster and that cluster has passed all the internal thresholds, the algorithm selects only the pixels belonging to this cluster and apply the Inverse Wavelet Transform. In this way Waveburst turns in the time domain only the part of the data stream referring to the reconstructed signal.

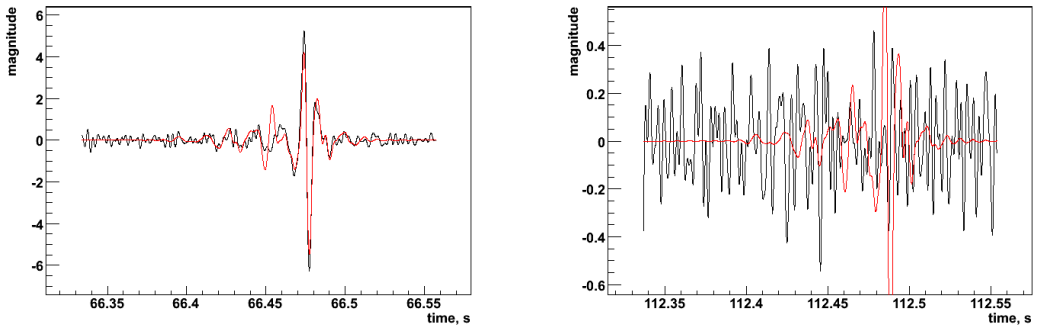

Figure 3.10: *Example of a reconstructed waveform for an event from a background analysis. The red line is the reconstructed signal, the black is the original data. Left: L1 detector, right: V1 detector*

Once obtained the original waveform, we can calculate all the related informations: frequency, time, etc... Waveburst defines the quantity of central frequency and central time, that are calcolated with a weighted mean of the times or frequencies associated to the pixels composing the cluster. The amplitude pixels are used as weights.

The **source coordinates reconstruction** is performed using the information from the likelihood calculated over the sky. The used principle is to define a rank on the sky positions and associate the reconstructed direction to the first positions given by the ranking. For each reconstructed event, given a sky position i, Waveburst ranking it defining the **Sky Statistic**:

$$S_i = \frac{L_i \cdot cc \cdot P}{E} \tag{3.62}$$

where L_i is the likelihood on the sky position i, cc the network correlation coefficient (Sec. 3.3.2), P the penalty (Sec. 3.3.1) and E the coherent energy of the reconstructed event.

The source coordinates reconstruction is not performing as well as desired, so there is a dedicated study on this aspect. We discuss in Sec. 4 in a detailed way what has been done to improve the reconstruction of source coordinates.

3.2.7 Calculation of observation time

The calculation of the observation time is not related to the event reconstruction, but are useful for other quantities, for instance the rate of events in the considered period.

Live time is determined by a coincident lock segment of all detectors in the network and by the data quality flags (DQ), which are list of time period when analysis is applied. Primary data segments (science mode segments with some DQ flags included) are used to generate a list of shorter segments in order to run the CWB jobs. Final data segments (with all DQ flags included) are used during the producton stage to set the time segments to analyse. Calculation of live time, especially for non-zero lag, is not so trivial, when final data quality is applied. Procedure can be simplified calculating live time during the production stage. From the list of final data segments, the pipeline defines a time-frequency array (veto$[i, j]$) where pixels are set equal to 0/1 if their time is outside/inside the data quality intervals. When the analysis of the LTF pixels is performed, they are accepted or rejected according to the array. Live time is defined as the ratio between accepted pixels and the total number of data samples multiplied for the duration of the data segment. Live time is stored in the output CWB file with the triggers produced by the job.

3.3 Post-production analysis

When the detector noise is Gaussian and stationary, the maximum likelihood is the only statistical tool required for detection and selection of the GW burst triggers. In this case the pipeline false alarm and false dismissal probabilites can be controlled by setting a threshold on likelihood at the output of the coherent event generator. Unfortunately, the detector outputs are contaminated by glitches and additional selection cuts need to be applied to distinguish glitches from GW signals. These selection cuts are an ad hoc procedure and depend on the network configuration and on the search parameters. Pipeline has therefore been divided into two stages: the coherent event generator and the post-production analysis. The coherent analysis is performed at the production stage and all the coherent statistics such as likelihood, null stream, etc, are stored in the output trigger files. The post-production stage deals with the selection of the optimal set of statistics and does not require to run the coherent algorithms.

Here we define the consistency cuts applied to the second part of the pipeline.

3.3.1 Penalty

In a previous paragraph we have introduced the energy disbalance for the single pixel (Eq. 3.44). We can generalize the definition for the reconstructed event including all the pixels belonging to the cluster:

$$\sum_k \sum_{i,j} (X_k[i,j]\xi_{\sigma k}[i,j] - \xi^2_{\sigma k}[i,j]) = 0 \tag{3.63}$$

where k is the detector index and i, j refer to the pixels. We know from likelihood analysis that this relation is equivalent to te orthogonality between detector response and null stream, i.e. $\sum_{i,j}(X[i,j] - \xi_\sigma[i,j]) \cdot \xi_\sigma[i,j] = 0$. However, this is a network quantity, that does not assure that the same equality is satysfied for each detector.

We should require the orthogonality between signal and noise separately for each detector. The reason for this is to prevent the reconstruction of un-phyisical detections when the response energy is greater than the total energy in the detector data stream, i.e. we require:

$$E_k > S_k \longrightarrow \begin{cases} E_k = \sum_{i,i} X^2_k[i,i] \\ S_k = \sum_{i,i} \xi^2_{\sigma k}[i,i] \end{cases} \tag{3.64}$$

These constraints are particularly important for all sky searches to find a correct source location and help to reduce the false alarm rate. For CWB these constraints

are applied in the form of a penalty factor, by penalyzing those points in the sky where the constraint is not satisfied.

$$P_f = \max_k P_k; \quad P_k = \begin{cases} \sqrt{\frac{E_k}{S_k}} & E_k < S_k \\ 1 & E_k > S_k \end{cases} \tag{3.65}$$

In the post-production it is required that $P_f > P_{th}$

3.3.2 Network correlation coefficient

The Likelihood is a quadratic form and we can write it in a matrix form.

$$\begin{aligned} L &= (X \cdot e'_+)^2 + (X \cdot e'_\times)^2 = \\ &= \sum_{mn} [(X_m e'_{+m} X_n e'_{+n}) + (X_m e'_{\times m} X_n e'_{\times n})] = \\ &= \sum_{mn} L_{mn} \end{aligned} \tag{3.66}$$

where the $'$ terms are defined in the DPF.

The sum of the off-diagonal terms is the coherent energy E_c.

$$E_c = \sum_{m \neq n} L_{mn} \tag{3.67}$$

We can costruct a correlation coefficient using coherent energy and the total reconstructed energy of noise in units of the noise RMS (N_{ull}):

$$N_{ull} = |X - \xi_\sigma|^2 \tag{3.68}$$

We define the network correlation coefficient as:

$$\mathrm{cc} = \frac{E_c}{E_c + N_{ull}} \tag{3.69}$$

By setting a threshold on the correlation coefficient *cc* one effectively compares the null energy with the coherent energy. This is a much safer selection cut than a null stream cut (where the null energy is compared with the estimated noise energy) because in any realistic data analysis there is always some residual energy left in the null stream. For strong gravitational waves the energy of the residual signal can be much larger than the noise energy resulting in the false rejection of the GW signal.

3.3.3 Effective Correlated SNR

As the network correlation coefficient, we can define Pearson's correlation coefficient:

$$r_{mn} = \frac{L_{mn}}{2\sqrt{L_{mm}L_{nn}}} \tag{3.70}$$

using these coefficients we construct the *reduced correlated energy*

$$e_c = \sum_{m \neq n} L_{mn}|r_{mn}| \tag{3.71}$$

We use this quantity to define *effective correlated SNR*

$$\rho = \sqrt{\frac{e_c}{N}\mathrm{cc}} \tag{3.72}$$

where N is the number of detector.

The Effective Correlated SNR is related to the Likelihood values, which is the sum of square SNR of single detectors (Eq. 3.1.2). Infact, if we consider a totally correlated signal ($cc = 1$) $e_c = L/2$, so $\rho = \sqrt{L/2N} = \sqrt{\sum_i SNR_i^2/2N}$. The second equality becomes from the consideration that for a totally correlated signal the sum of off-diagonal elements of likelihood matrix is equal to the the diagonal. The coherent energy (off-diagonal terms) cannot be larger than the uncorrelated energy (diagonal terms), this would be an unphysical situation. The maximum value of the uncorrelated energy is equal to correlated one, and it is reached when the signal is completelly coherent, conseguently the e_c values is equal to half of likelihood value. If the signal is not completelely coherent it is satisfied by the following disuguagliance: $\rho < \sqrt{\sum_i SNR_i^2/2N}$. So, assuming that for the most part of the signals the network correlation is near to one, and $\rho \approx \sqrt{\sum_i SNR_i^2/2N}$, the Effective Correlated SNR is proportional to the network SNR.

3.3.4 Effects of post-production cuts

The effective correlated SNR and the network correlation coefficient are very useful to charcterize the events reconstructed from background or simulated events. We expect that noise excesses should not be correlated, so the *cc* value should be small. Where are GW should be very correlated, and so their *cc* values should be near to one.

The value of effective correlated SNR depends strongly on the energy of the signal, so its value could be great both for energetic GW and for great noise excesses. However, we expect that glitches with large value of ρ tipically have low value of the correlation coefficient. Morevoer, background triggers with large

values of *cc* are usually produced by casual fluctuations of the detector noise which are not energetic, so they can be effectively removed by setting a threshold on ρ (Fig. 3.11).

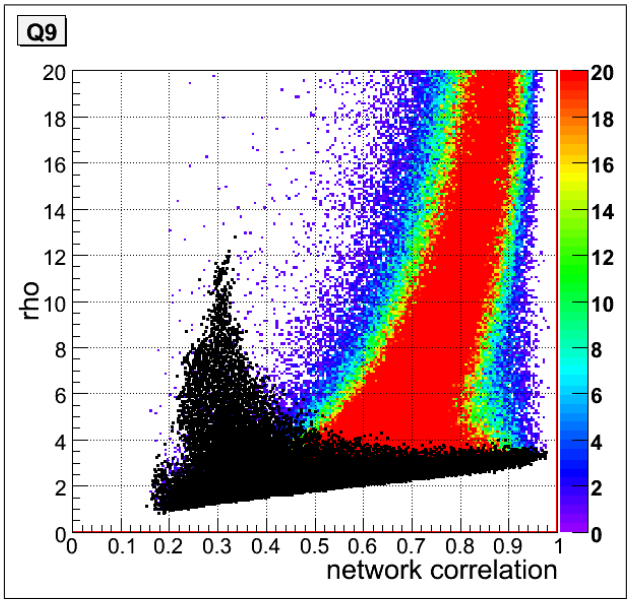

Figure 3.11: *Distribution of ρ and cc quantities for background events (black) and Sine Gaussian simulations (coloured dots). Network is composed of the detectors: L1, H1, H2, V1.*

3.4 Coherent Event Display

The list of triggers resulting from a certain analysis can be used to perform interesting checks regarding the noise detector or the algorithm performances. After a background analysis, for instance, it is a good idea to study the characteristics of the loudest reconstructed events, to understand their origin. If these events have high SNR, this could be due to particularly noisy data stream, or to undiscovered bugs on the algorithm, or simply to the statistics. So, it is useful to have complete informations about these events.

However, the informations stored in the merger ROOT files are not enough to perform this depth study, since they are not present useful informations like the comparison between the original and the reconstructed h(t), the Time-Frequency map, the likelihood skymap, the sky statistics, and so on... Otherwise it is not necessary to keep these informations for all the events belonging to the trigger lists (it would be too computationally expensive).

To resolve these issues it has been developed an useful tool which products and saves additional informations about the reconstructed events. This instrument is called **Coherent Event Display** (CED) [70] and it belongs to Waveburst libraries.

The CED is enabled by a designed bool variable in the parameter file and creates for each reconstructed event a directory where it saves esplicative figures containing the informations listed in the following sections. A complete description of all the reconstructed parameters can be found in the CED Technical Documentation [71], in the following sections we list briefly what they are.

Jobs and Event parameters

This first section details the parameters which are specific for the job and for the event we are interested for:

- Event times: start, stop and central GPS, recorded along with the central time relative to the start of the data job used;
- h_{rss} which represents the reconstructed strain;
- Likelihood (Eq. 3.10);
- Network correlation (Eq. 3.69);
- Reconstructed sky locations (θ, ϕ). The coordinate system is the Earth fixed frame with $\phi = 0$ corresponding to the Greenwich Meridian, $\theta = 90$ corresponding to the equator;
- Size: the number of pixels in the appropriate TF decomposition level that make up the detected event;
- Frequency informations: the low and high frequency limits of the events, the central frequency as wighted by the likelihood, the total bandwidth and the wavelet decomposition requency.

Network Data Matrix

This section contains what we call Network Data Matrix (Eq. 3.66), which represents the total detected energy of the event. The diagonal components of the matrix constribute to the incoherent energy while the off-diagonal components to the coherent energy.

Time-Frequency Maps

This section shows the wavelet coefficients normalised by the noise RMS as a function of time and frequency for each detector.

Likelihood Time-Frequency Maps

This section contains the Likelihood (Eq. 3.10) as a function of time and frequency for the whole network. There are three different plotting styles for this figure: Pixel, Shaded and Cluster. The *Pixel* shows all the pixels in a time-frequency range near the event; the *Shaded* interpolates between adjacent data points to give a smoother appearance to the map, and the *Cluster* shows only the event cluster requested.

Reconstructed Detector Response

This section contains three tpyes of plots: one of this shows the reconstructed strain for the detected event, while the others contain the comparison between two different reconstructed waveforms, the reconstructed strain and the band limited detector output. The difference between these two plots is that the former use only the reconstructed signal and the latter uses both the signal and the noise to make this comparison.

Sky Maps

Several reconstructed parameters have a dependency on the sky locations, so this section shows a series of skymaps with the values of the parameter as a function of the sky location. The list of such figures are:

- Network Sensitivity to plus and cross polarization, defined in the Dominant Polarization Frame (Eq. 3.17);
- Penalty factor (Sec. 3.3.1);
- Network correlation (Eq. 3.69);
- Likelihood (Eq. 3.10);
- Sky Statistics (Eq. 3.62);
- Coherent Energy, as the sum of off-diagonal elements of Network Data Matrix (Eq. 3.66);
- Null stream (Eq. 3.68);

- Total Energy Disbalance (Sec. 3.1.5):

$$S = \sum_k |\Delta_k| \tag{3.73}$$

- Disbalance Asymmetry:

$$\frac{L_m}{L_M} \quad \text{where:} \begin{cases} L_m = \sum_k \min(X_k \xi_{\sigma k},\ \xi^2_{\sigma k}) \\ L_M = \sum_k \max(X_k \xi_{\sigma k},\ \xi^2_{\sigma k}) \end{cases} \tag{3.74}$$

The two quantities L_m and L_M we have just introduced are related to the maximum likelihood L_{max} (Eq. 3.43) and the total energy disbalance S. Infact we can write:

$$L_{max} = \sum X_k \xi_{\sigma k} = \sum \xi^2_{\sigma k} = (L_m + L_M)/2 \tag{3.75}$$

and

$$S = \sum_k |\Delta_k| = \sum |X_k \xi_{\sigma k} - \xi^2_{\sigma k}| = L_M - L_m \tag{3.76}$$

Finally we can write:

$$\begin{cases} L_m = \frac{2L-S}{2} \\ L_M = \frac{2L+S}{2} \end{cases} \tag{3.77}$$

and the Disbalance Asimmetry:

$$\frac{L_m}{L_M} = \frac{2L - S}{2L + S} \tag{3.78}$$

Chapter 4

Waveform parameters reconstruction

In the present chapter we focus on the problem of reconstructing the waveform parameters and the source sky coordinates. This is an extremely challenging purpose in the GW data analysis because the knowledge about the source and waveform physical characteristics enhances the detection confidence of the signal and offers the chance to interact with other observational experiments that could produce an optical or electromagnetic (EM) counterpart of the particular event at the selected sky location.

In fact, we would generically expect some release of energy in the EM spectrum for many gravitationally detectable astrophysical processes (but not all). So, it is plausible to expect a detectable EM transient to accompany a detectable GW burst. Moreover, the detection of EM emission is easier than the GW detection. Some dual GW/EM source mechanism that we can expect are [72]:

- **Merger** of double neutron stars of black hole/neutron star. Models and numerical simulations [73, 74, 75] have speculated the ejection of neutron rich matter from such processes which would then decay, fueling a nuclear fireball. However, there is no empirical evidence on this: the above models are based on expected radiactive decay of ejecta from mergers.

- **Supernovae**. Comparison of expected GW emission and observed supernova EM emission shows that taking in account the predicted detector sensitivity from mid-2009 any supernova detectable in GW would be quite bright in the EM spectrum [76].

- **Short Gamma Ray Bursts (GRB)** [77] are of particular interest to gravitational wave physics. These events are believed to be associated with double neutron star mergers, an so would emit significant gravitational radiation

[78].

The potential benefit of coupling EM observations with GW is the opportunity to link a GW event candidate with a high-energy astrophysiscal event, thus promoting an otherwise speculative candidate to a true detection. This may be considered as an additional coincidence test. A recent estimate [79] found that this approach would have doubled the volume of space searched in joint period from 2006 to 2007 of the Ligo-Virgo collaboration. Moreover, from mid-2009, when the expected sensitivity is significantly increased [80], such efforts are critical to improve the effective reach.

In the case of a confirmed gravitational wave burst, the extractable information increases dramatically if the burst is coupled with an EM counterpart. Modeling an astrophysical source mechanism is best accomplished with the aid of "multi-messenger" astronomy. The rich possibilities of multi-messenger astronomy in the context of gravitational waves has attracted much attention lately [81] and was the theme of the 2008 and 2009 Gravitational Wave Data Analysis Workshop [82, 83]. The EM counterpart to a GW signal could provide from the sky position the host galaxy type and the distance of the source; moreover, the emission characteristics (light curve, spectrum, ...) could provide additional clues to the underlying source mechanism.

LSC collaboration uses regularly GRB and Soft Gamma Repeaters (SGR) to perform sensitive searches for GW Bursts [84, 85, 86]. These searches (called **ExTrig**) have been highly successful in demonstrating the power of using EM coincidence to improve GW sensitivity. However ExTrig searches demand a detectable gamma-ray signal and there is no reason to expect that all GW/EM sources emit in gamma-rays.

The plan of the LIGO-Virgo collaboration is to implement an **on-line** analysis [87]: data from the three site network is analyzed in pace with real time, obtaining significant triggers with a time-lag of 10-30 minutes. The reconstruction of triggers includes an estimation of source position in the sky, which includes the most likely position and an uncertainty region. Optical followup observations are initiated for event candidates which are sufficiently well localized, in the field of view of a suitable telescope, and above some significant thresholds, which are adjusted to limit the rate of candidate events. Then the contacted telescopes would capture images at sky positions based on the GW trigger informations.

The choice of partner telescopes is affected by the expected reconstruction perfomances of the GW detectors, which is at the level of a few degrees [88]. This suggests to seek partner telescopes with comparably large fields of view (FOW). Further, we wish to perform many short observations at times which cannot be specified in advance, so automated or semi-automated telescopes are desiderable, preferably those that are designed with responding to triggers in mind. Obviously

Name	FOV (degrees)	Aperture (m)
Swift	0.3x0.3	2.7
ROTSE III	1.85x1.85	0.45
TAROT	1.85x1.85	0.25
SkyMapper	2.4x2.4	1.3
QUEST	4.5x3.5	2.2
Palomar Transient Factory	3.4x2.3	1.2
Pi of the Sky (prototype)	20x20	0.07 (effective)
Pi of the sky (final)	20x20 or 40x40	0.28 or 0.14 (effective)

Table 4.1: *Table of wide-filed optical telescopes with few of degrees (FOV) of at least a few square degrees.*

are preferred telescopes with the capability to image any region of the sky at any time. Some telescopes which satisfies these requirements are: ROTSE [89], TAROT [90], SkyMapper [91], Palomar Transient Factory [92], QUEST [93] and Pi of the Sky [94]. Their main characteristics are reported in the Tab 4.1. Others interesting telescopes could be: LONEOS [95] and Zadko [96], the last one is dedicated to transient searches and collaborates with TAROT.

An interesting telescope is the Swift Gamma Ray Burst Explorer (or simply Swift [97]), which is capable of performing multi-wavelength observations of GWBs and their aftermath. It is composed of three different telescopes: BAT which triggers the events and calculates the direction coordinates; XRT, that detects the X-ray emission; and UVOT, which studies the emission in the visible and ultra-violet frequency band. Involving this telescope, there are three targets of opportunity observations granted for the cycle of detection in the time period between April 2009 and March 2010.

As we have just esplained in Sec. 3.2.6, waveform parameters reconstruction can be described by two principal aspects: the waveform and source directions. In this section we will focus on the modifications made on CWB algorithm made to improve the performances related to these aspects. This is an **on-going** work, so the presented results are not final, and the group is continously working on it to obtain the desired goal. This could include also strong modifications on Waveburst that would change in a significant way the algorithm implementation respect to the one reported in the Sec. 3.

In particular, regarding the source direction reconstruction performances the LIGO-Virgo collaboration has started a project: the **Position Reconstruction Challenge** (**PRC**) which has the goal to give a reliable reconstruction of source sky position. So, in this chapter we will explain the PRC framework which goals we have just expressed above.

In this chapter we test the algorithm performances on two waveform type: Sine Gaussian and White Noise Bursts. We postpone to a following chapter (5.2) the characteristics of these waveforms and the motivations behind this choice. We briefly explain now their main characteristics.

Sine Gaussians (SG) are linear or circular polarization waveforms described by an equation of the type:

$$h = e^{-(t/\tau)^2} Sin(2\pi f t) \tag{4.1}$$

where the decay time is related to the central frequency: $\tau = \frac{Q}{\sqrt{2}\pi f}$. We usually describe a particular Sine Gaussian giving the values of frequency f and Q. In this chapter we test performances on SGQ9 and SGQ3 with linear polarization, and SGQ9 with circular polarization (specified with the label "C" in the name).

White Noise Bursts (WNB) are composed of random polarization white noise which are restricted in defined frequency band and time period. To describe them, we list the starting frequency, the frequency band and the duration.

Before introducing the PRC project, which is the main subject of this chapter, we briefly describe some issues about the waveform reconstruction.

4.1 Waveform reconstruction

We have introduced in Sec. 3.2.6 how the algorithm reconstructs the waveform: it applyes the inverse wavelet transform to the cluster pixels obtaining the signal in the time domain. An example is reported in Fig. 4.1, where the red line is the injected signal, and the blue line is the reconstructed one.

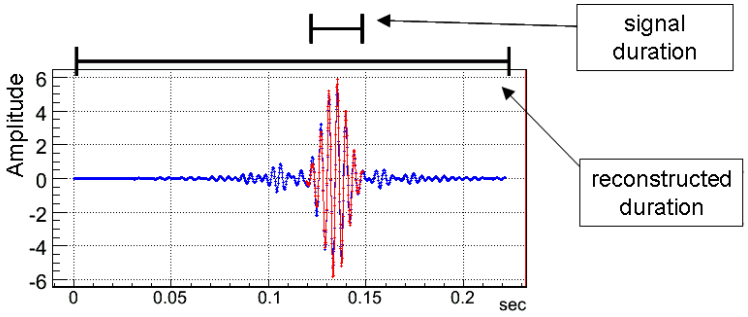

Figure 4.1: *Example of waveform reconstruction applied on a Sine Gaussian signal with $f = 235$ Hz and $Q = 9$. Red curve is the injected signal, and blue curve is the reconstructed one.*

We can see that reconstructed signal shows noise tails outside the duration of the original signal. In principle these tails could affect the likelihood calculation,

introducing a small bias on the reconstruction of source direction. However, these noise tails carry a small percentage of the total signal energy, so we do not expect great biases.

To characterize the waveform reconstruction performances we have introduced a new quantity, N_{orm} which is defined taking in account the difference between the amplitudes of reconstructed and injected waveforms:

$$N_{orm} = \frac{\sum_n (Rec_n - Inj_n)^2}{\sum_n (Inj_n)^2} \tag{4.2}$$

where Rec_n and Inj_n are respectively the reconstructed and injected sampled amplitude, and n refers to the sum on the samples. This quantity is a sort of percentage error on the reconstruction.

Results are shown in Fig. 4.2 where we report the median value of N_{orm} as a function of the network SNR. We can see that the value of N_{orm} decreases when SNR increases, as we expect, because for higher SNRs the signal is better localized in the TF map and so reconstruction is better performed. However, for SNR greater than a great value (~ 100) the N_{orm} becomes almost constant. This is due to the contribute of the noise tails.

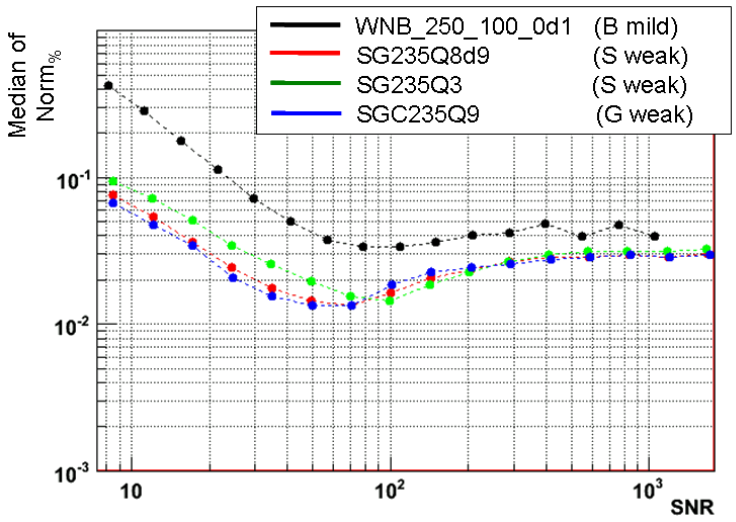

Figure 4.2: *N_{orm} as a function of SNR for some waveforms with characteristic frequency near 200 Hz. Applied constrains are listed in the figure.*

The effect of the noise tails can be estimated calculating the N_{orm} quantity discarding about the amplitude pixels outside the signal duration. Referring to Fig. 4.1, this means to consider only the times belonging the red curve, which we

call *Bulk Signal*, with respect to the blue curve, called *Full Signal*. Fig. 4.3 shows that N_{orm} values for Bulk Signals is less than Full Signal ones. This difference is more significant for high SNRs, due to the fact that noise tails contribute more for this case. However, the N_{orm} constant behaviour for high SNR is visible also for Bulk Signal, due to the fact that reconstruction is affected by the unavoidable detector noise.

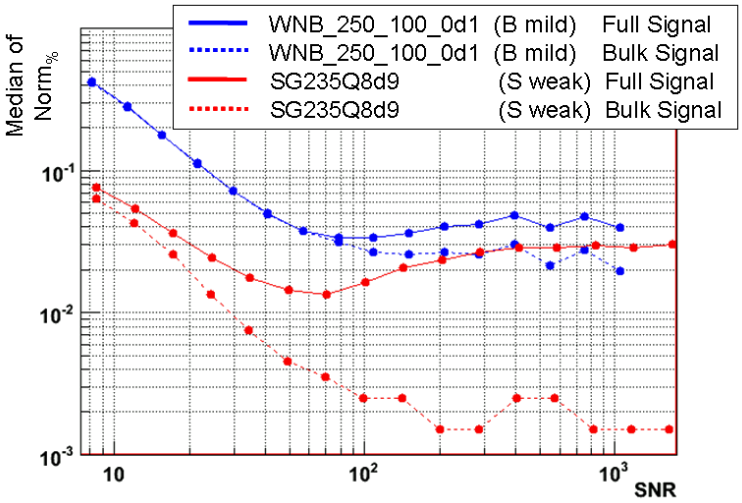

Figure 4.3: *Comparison of N_{orm} for Full (continous line) and Bulk (dashed line) as a function of SNR for some waveforms with characteristic frequency near 200 Hz. Applied constrains are listed in the figure.*

4.2 PRC framework

A single GW detector is not able to reconstruct the coordinate of the source generating the GW that it has eventually detected. We need a network of more detectors to implement a procedure of source coordinates reconstruction. One example is based on triangulation [98] which calculats the source coordinates from time delays between the times of detection of the involved detectors. We explain now the implementation of this method.

Using the delay time between two detectors we can reconstruct a circle of possible source positions on the sky, centered on the ideal line joining the two detectors sites. Adding another detector, we can reconstruct three circles, one for each couple. In this case, the intersection points of these circles correspond to the source location and to its so-called mirror position, i.e. there is still the ambiguity

left (Fig. 4.4). There are some possibilities to remove this ambiguity, one of this is to add a fourth detector to the network (outside of the plane defined the others). Another possibility is to use informations from antenna patterns to relate them to the characteristics of the signal in the three detectors.

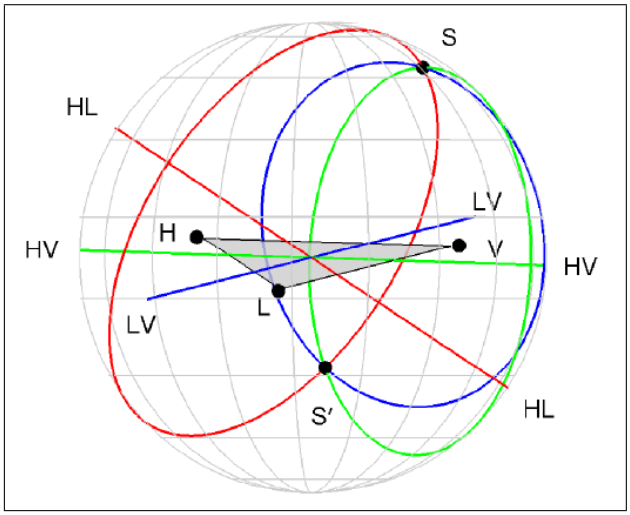

Figure 4.4: *Geometry of the network composed of the three LIGO-Virgo detectors (H, L, V). The locus of constant time delay between two detectors form a ring on the sky concentric about the baseline between two sites. For three detectors, these rings normally intersect in two locations: the source one (S) and the mirror image (S'). For four or more detectors there is an unique intersection points of all the rings.*

However, all of these approaches are affected by imperfect estimation of the signal characteristics, due to the noise. The triangulation method supposes to construct the time passage of the wave in each detector with a great precision. The other approach supposes to have a good estimation of the signal characteristics.

Coherent Waveburst approaches both these methods: likelihood, in fact, is calculated using the detectors antenna patterns and over the entire sky, which is equivalent to take in account the time differences between involved detectors.

In the following sections we explain what are the main characteristics of the PRC project.

4.2.1 Data Set

The PRC has been extensively tested on a small data set including 4 days of data jointly collected during May 2007 by LIGO and Virgo and called Joint Week 1

(JW1). However, the project is directed to a following period, started on mid-2009 and called S6/VSR2, where it is predicted that Virgo sensitivity should be comparable to the LSC detectors. So, some tests have been performed simulating 4 days using the same simulated sentitivity for all the detectors according the predicted S6/VSR2 sensitivity. Because of internal agreement of the LIGO-Virgo collaboration, in this thesis we will report only the results on **simulated data**.

4.2.2 Network

The network contains the three LSC-Virgo sites: L1, H1, V1. We do not include H2 detector because we do not expect significant improvements adding this detectors. Indeed H1 and H2 are co-located, so adding H2 on the network could improve glitches rejection, but not the coordinate reconstruction.

4.2.3 Simulated signals

We used simulated waveforms to test the efficiency of reconstruction. These waveforms belong to four different set, which we call **MDC**. We explain in Sec. 5.2 in details the characteristics of these waveforms and the motivations regards the choice. Each waveform is characterized by its frequency, and it shape. Three of the four sets are composed of Sine Gaussian waveforms, linearly or circularly polarized, the other is composed of randomly polarized White Noise Bursts.

In the Tab. 4.2 we list the waveform characteristics for each set.

Name	**Polarization**	**Low Frequency**	**High Frequency**
SGQ9	Linear	SG235Q9	SG1053Q9
SGCQ9	Circular	SGC235Q9	SGC1053Q9
SGQ3	Linear	SG235Q3	SG1053Q3
WNB	Random	WNB_250_100_0d1	WNB_1000_100_0d1

Table 4.2: *Waveforms used in the PRC framework for efficiency tests. For the explanation of the name label see Sec. 5.2*

We inject signals with hrss amplitude at Earth rescaled with respect to the reference value $2.5 \cdot 10^{-21}$ $1/\sqrt{Hz}$. In particular we use 14 rescaled amplitudes spaced in multiples of $\sqrt{2}$ where we expect steep change, and a factor of 2 spacing at the ends to increase the range. We inject signals in 46 different directions, roughly uniformly distributed in the sky (Fig. 4.5). The livetime of the period analysed is about 4 days and injections are made every 30 seconds. So, for each set of waveforms and each of the 46 sky positions, we have about 250 injections during the considered period for each injected amplitude factor.

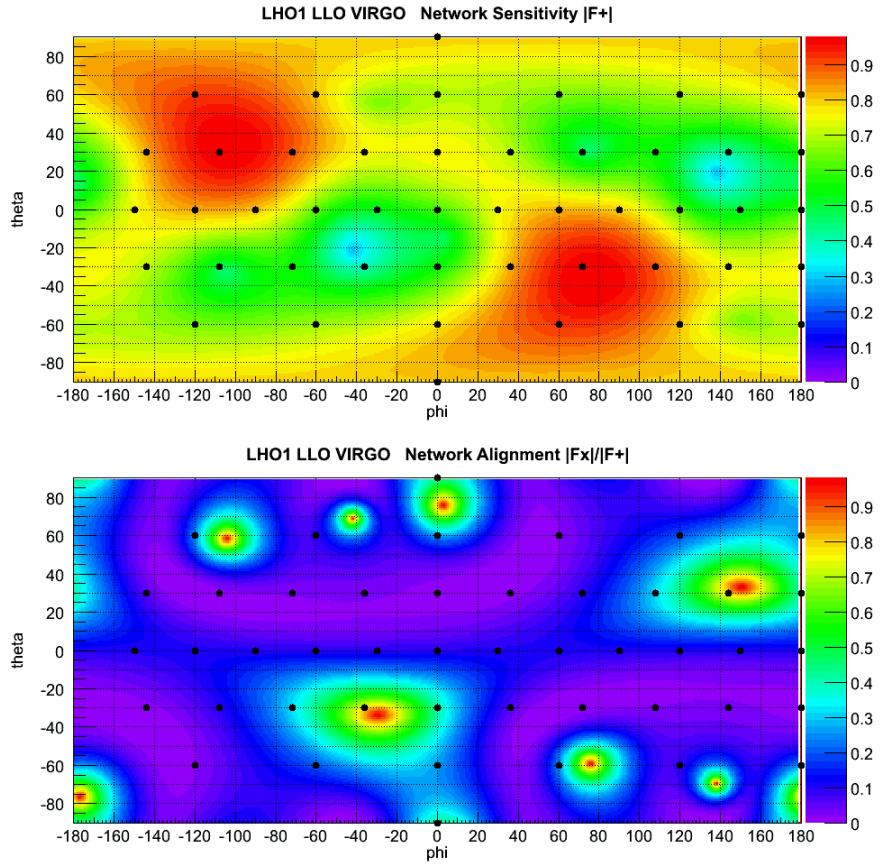

Figure 4.5: *The 46 sky positions chosen for the tests of PRC. Coloured background is $|F_+|$ (top) and $|F_+|/|F_\times|$ (bottom) values over the sky, where $F_{+,\times}$ are network antenna pattern.*

Moreover, we realised that this chosen set of directions is well representing the entire sky. In fact, we performed analyses considering sets of waveformes uniformly distributed in the sky, verifying that on average the reconstruction performances are the same for the two cases.

4.2.4 Algorithms

Three **Algorithms** are involved in the PRC projects: Coherent Waveburst [57], Triangulation [61] and Bayesian [100]. In this thesis we focus only on the first one discarding on the performances of the others. We are not interested in this work to compare results from the different pipelines or find out what is the best performing.

4.2.5 Reconstruction quantities

The PRC group has introduced the following quantities to characterize the performances and improvements of the algorithms. The first is the so called **Effective Error Angle**, which indicates the reconstruction error between the injection and the reconstructed directions. Obviously, more this quantity is smaller, more the algorithms reconstruction perfomances are better. A way to define this quantity is to take the solid angular difference between these two sky positions. However, not always the algorithm is able to resolve the ambiguity between source and mirror position, so this quantity could assume meaningless values. A more reliable way to define Error Angle is to take from the Sky Statistic (Eq. 3.62) the integrated area of the pixels which have likelihood value greater than the injection one. In other words, we rank the sky positions according to the corresponsing value of Sky Statistic, we define Error Angle as the area including pixels with Sky Statistic value greater than the injected (Fig. 4.6). With this definition this quantity can be considered as the minimal area in which we are sure to include the injection direction according to the established ranking. Obviously, Error Angle strictly depends on the tested waveform, because it could be defined only using simulated signal. It would be desiderable to have a quantity that has a general definition, i.e. is not dependent of the tested waveforms.

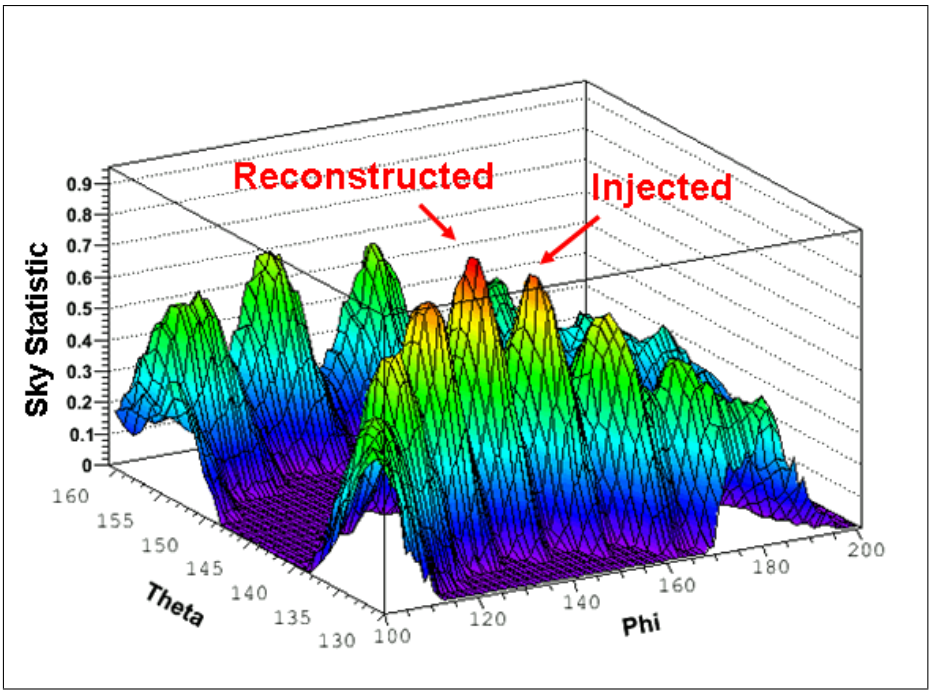

Figure 4.6: *Sky Statistic with definition of Error Angle.*

The PRC group has introduced the idea of a new quantity to resolve this issue. This is called **Error Region**: defined as an area in the sky that has a certain probability to contain the right direction. This approach could resolve also the ambiguity problem: Error Region could be a not connected region, so it could contains both the injection and the mirror positions when the algorithm is not able to resolve the ambiguity.

Each algorithm has a different method to calculate the Error Region, for Coherent Waveburst it is calculated converting the Sky Statistic to a **Probability Sky Map**, so to have a definition which is not dependent on the tested waveform. However, we never forget that the results performances of this quantity depends on the tested waveforms, and we could not generalize them to all the possible GWs.

Sky Statistic is trasformed to Probability introducing a **conversion** function. The conversion function is defined empyirically from theoretical considerations based on the simulations tests we have performed during this work. This means that the definition explained here would be probably changed in the future, according to the reconstruction performances, or to theoretical motivations which could lead to an alternative way to define the probability.

The conversion function associates to each sky pixel i a value depending on the Likelihood and other quantities related to the entire sky. For each pixel i we introduce the exponential function $P_i = e^{-aN_i}$ where the N_i term is the *null stream* of the reconstructed signal, which can be calculated as the difference between coherent energy and likelihood.

$$N_i = E - L_i \approx E(1 - S_i) \tag{4.3}$$

where we have substituted to the Likelihood the Sky Statistic, considering that $S_i \approx L_i/E$. The idea is that the minimal null stream associates the maximum probability. The a term is introduced to normalize the probability ($\sum_i P_i = 1$). Simulations tests suggest to introduce the null stream variance N_0 in this way:

$$P_i = e^{-a(N_i/N_0)} \tag{4.4}$$

where the null variance is defined as:

$$N_0 = n(K + K_0) + pr \cdot E \tag{4.5}$$

where K is the cluster size, n is the number of detectors, K_0 and pr are two parameters. The idea is that the null variance can be calculated as the sum of variance of the single pixels componing the cluster. Since pixels amplitude are normalized for the noise, the variance of the single pixel is equal to one, so the total variance is simply the sum of the total pixel in the cluster multiplied

by the number of involved detectors. The introduction of K_0 and pr are due to algorithm approximations, which we estimate as a constant term (K_0) and a energy depending term ($pr \cdot E$). There is no theorically way to select K_0 and pr, so their values are chosen according to test results.

Once obtained this Probability Sky Map, Error Regions are defined as the area containing the integrated probability. Error Region X% is the area in which the *integrated probability* is X% (Fig. 4.7).

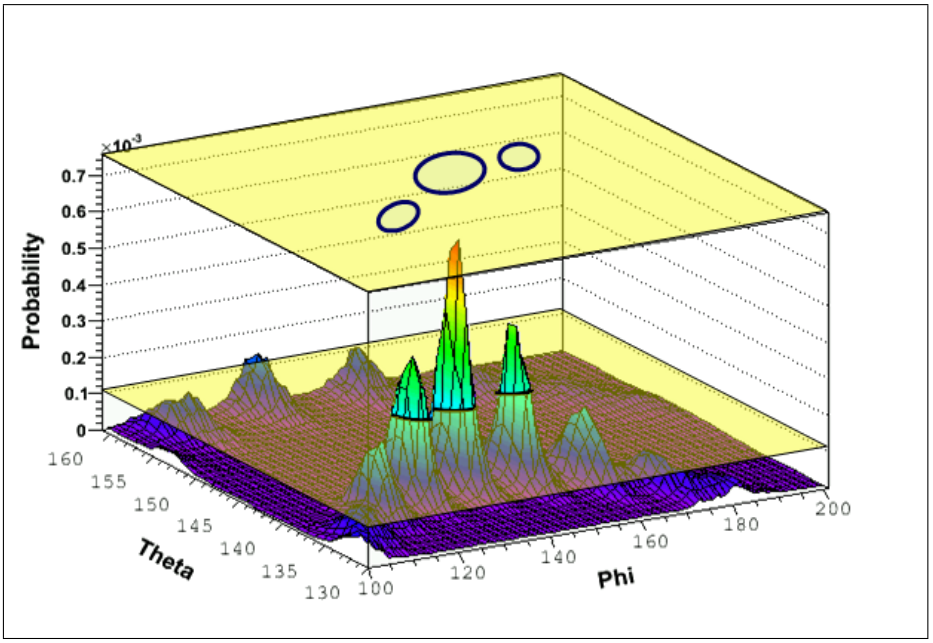

Figure 4.7: *Probability Sky Map with definition of Error Region.*

4.3 Technical Issues

Following the PRC project characteristics, the pipeline has been modified to improve direction reconstruction without reducing the efficiency detection of the starting version. In particular, the algorithm has been mainly modified in two parts: where likelihood is calculated, and what is regarding the sky representation. Moreover, Error Regions calculation has been implemented in the code.

In this section we deal about some of these innovations and the problems or issues related to them.

4.3.1 How Regulators and Energy Disbalance affect Reconstruction

We have introduced the Chapter 3.1.4 the definitions of Regulators used by CWB, explaining how the performances change according to the regulator chosen, considering the fact that if we introduce more assumptions we increase glitches rejection but we decrease detection efficiency and we do not want to penalize too much the detection efficiency.

Obviously, the best reconstruction is performed by the weak regulator, which gives also the greatest number of background glitches. We know that a certain pipeline is preferable when it has higher efficiency for lower glitches number. The regulator which shows the smallest number of glitches is the *hard* one, because it consider only signals with plus polarization. This conditions is satisfyed for some positions in the sky, but it is sfavourable for others. However, the hard regulator is a good choice for a network of two detectors, instead for three detectors the efficiency is too much penalized.

A good compromise seems the *mild* regulator, which is ongly dependent on antenna patterns, and so it is strictly affected by a great probability to describe the direction of a GW signal.

However, only simulation tests can give us the information of what is the optimal choice of regulator to have a good reconstruction and efficiency and a low false alarm rate.

4.3.2 How Sky discretization affects Reconstruction

Coherent Waveburst calculates likelihood functional over all the sky locations to find the source direction and also to construct sky probability which is used to define Error Regions. This procedure is made on a finite number of points in the sky. The set of these points is related to **sky discretization**.

A point of the sky is labelled by **polar coordinates** θ and ϕ. In this thesis we adopt the Earth fixed coordinate system, where $\theta = -90°$ refers to South Pole, $\theta = 0°$ is the equator line and $\theta = 90°$ is the North Pole. The angle ϕ goes from 0 to 360°, the 180° angle corresponds to the Greenwich meridian. Selecting an arbitrary angle resolution, we can discretize sky positions using step values of polar angles. For, instance, if we consider an angular resolution of 0.5×0.5, the sky discretization is composed of (180 / 0.5) $\times$ (360 / 0.5) points. The same principle is adopted for all the possible angular resolutions. Obviously, the chosen resolution defines what is the minimal error on the coordinate estimation.

However, in this way the angular resolution is not constant for all sky positions, because polar angles are defined on a spherical surface. This is the same situations

of meridian and parallel length, which decreases approaching the pole. The idea is to change nominal ϕ resolution varying θ angle to mantain the same effective resolution. So, if r is the Earth radius, the length of the parallel at latitude θ is $2\pi r cos(\theta)$. So, the mantain the same effective resolution, it is necessary to define a nominal resolution of the type: $(\Delta\theta, \Delta\phi) = (0.5, 0.5/cos(\theta))$. So, the number of points considered for each theta is equal to the near integer of $360/0.5 \times cos(\theta)$. With this definition, considering a nominal resolution of (0.5×0.5) degrees, the number of considered directions in the sky is equal to 165013.

This is the resolution we adopt for this thesis. Likelihood and sky statistics is calculated for all of these sky positions. We know, therefore, that the likelihood is calculated over the sky, using time delays on the data (Sec. 3.2.5), which are discretised. For instance, considering a sample rate equal to 16384 Hz, the time delay discretization is $\approx 63\mu$s. We can construct a correspondence between time delays and polar angles, using the triangulation principle, so that each sky position can be defined time delays[1]. Discretizing time delays, we discretize sky positions. So, using two time delays we have a number of discretized points equal to the square of the sample rate used, but the discretization is not uniform in the sky as the previous one, because the correspondence between time delays and polar angles is not linear.

This creates a problem, because for each position given by the angular discretisation Waveburst calculates the likelihood using the time delays couple which identifies the nearest sky position. However, for some region in the sky (especially near the detector plane) time delay discretization is less performing, so more points from angular set are associated to the same point from delay discretization (Fig. 4.8). This can create problems on direction reconstruction.

4.4 Results

The Coherent Waveburst algorithm has been modified significantly during the development of PRC project. As usually, an important modification is referred by a change of version. When the project starts, the CWB version was the one used for S5/VSR1 analysis (wat-4.7.3), the mean changes performed from that version regard about the introduction of new regulators (Section 3.1.4) and constrains (Section 3.1.6) and about the definition and implementation of Error Regions (Section 4.2.5). In this thesis we report results using the Wavebrust version **wat-5.0.9**.

[1]This is not a bi-univoc transformation, because we know that a couple of time delays correspond to two different positions in the sky, it is the ambiguity problem.

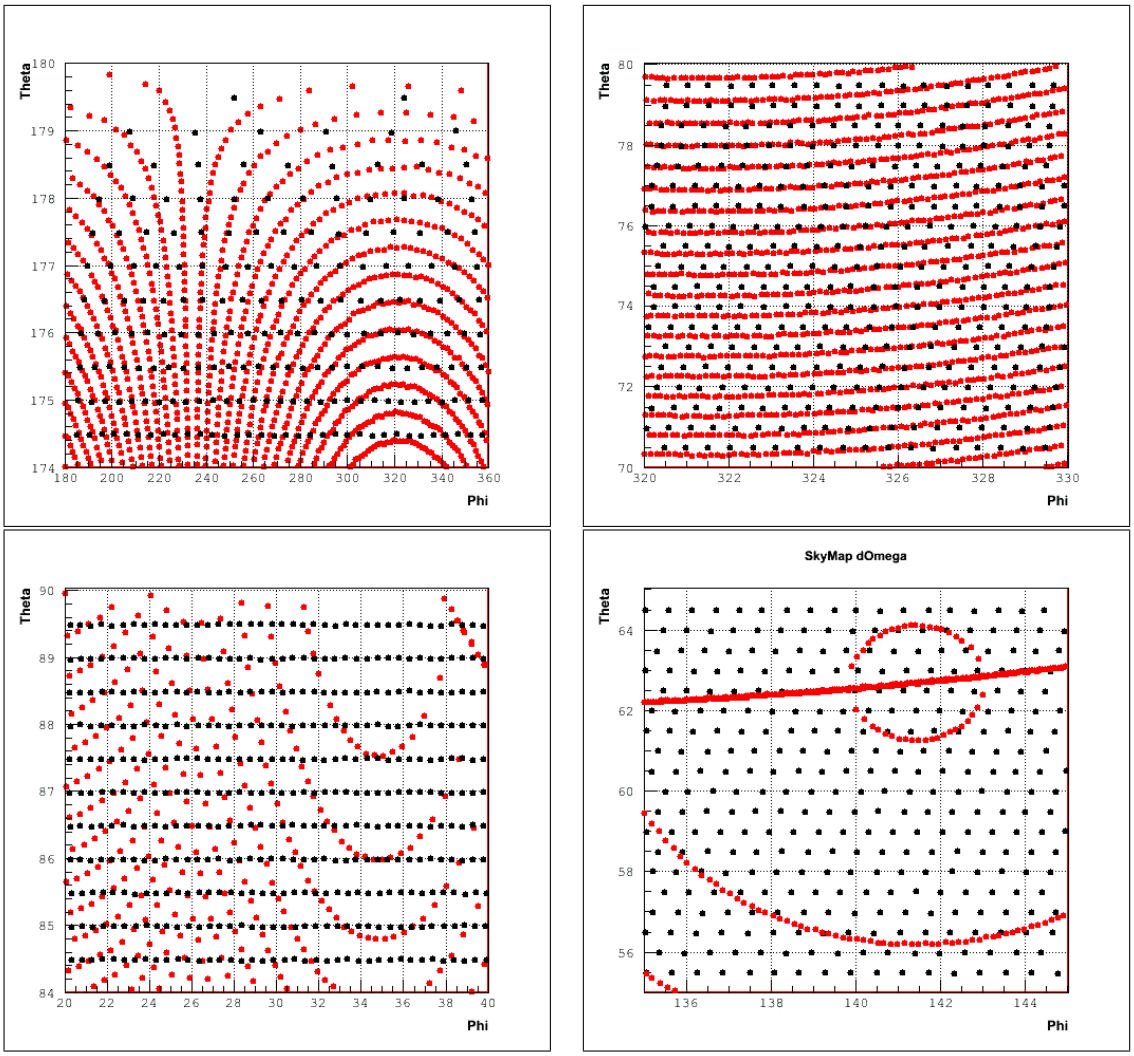

Figure 4.8: *Black points: sky discretization using a nominal angular resolution of 0.5×0.5 degrees. Red Points: sky positions corresponding to time delays discretizations assuming a sample rate equal to 8192 Hz. The figures show different sky regions: far from the detector plane (top) the red points are better associated to red ones respect when we are near the detector plane (bottom).*

We decide to use for each MDC set the **mild regulator** and the **unmodelled constrain**, this because this situation is adopted by the on-line analysis. However, for Sine Gaussian waveforms we implement test also using **linear**, **circular** or **elliptical** constrains, to test their effects on reconstruction. In these cases, moreover, we adopted the **weak** regulator because the adopted constrain is enough selective on the ellipticity to have a good rejection of glitches. We expect

that these tests show better results than for unmodelled constrain, and we call the **Best** choice when we adapt the constrain to the polarization of the injected signal (Tab 4.3).

MDC Type	**Best**	
	Constrain	**Regulator**
WNB	unmodelled (B)	mild
SGQ9	linear (S)	weak
SGQ3	linear (S)	weak
SGCQ9	circular (G)	weak

Table 4.3: *Constrains and Regulators for the **Best** choice adapted to the MDC type.*

Effective Error Angle

We report Effective Error Angle results using the following Figures of Merit (FOM)

- **Cumulative curves**
  Cumulative curves show for each point the percentage of events which have a certain value smaller than the one reported in the x-axis.

- **Sky Maps**
  These figures show the behaviour of a certain quantity related for each for the 46 injections directions. On the axis there are θ and ϕ values, and for each directions a number or a circle (which width is proportional to the value of the quantity) indicates the related value. Instead of the previous figures, which show the general behaviour of the involved quantities, these FOM show characteristic related to the injection directions and can be used to understand or verify problems related to some particular sky directions.

- **SNR dependence**
  These figures show the median value of Effective Error Angle as a function of Energy. From this plot we have more informations: the median value at several energies (especially if we are interested in a particular energy), the asymptotic value of these quantities (and at what energy it is reached), and so on. So, this figure is useful to verify more aspects regarding the performances of the pipeline.
  We use an empyrical equation to fit the behaviour of Effective Error Angle vs Energy:

$$\text{Median} = A + \frac{B^2}{SNR^2} \tag{4.6}$$

We report in Figs. 4.9 and 4.10 cumulative curves for all the waveforms. We use the Bmild search for WNB, and the Iweak search for SG. The green line is located at y-axis value equal to 50%, i.e. the median value. The Median value is reported in the legend for each waveform. The step on 0.5 degrees is related to the sky angular resolution used in the analysis.

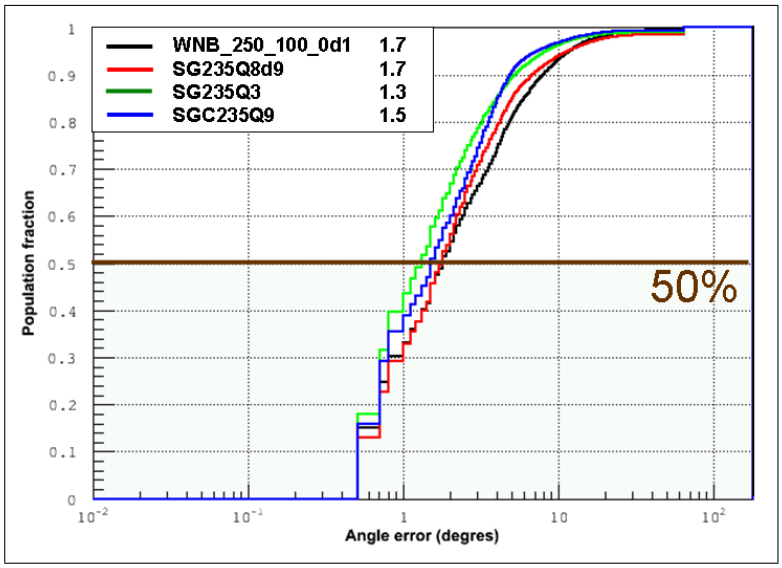

Figure 4.9: *Cumulative curves of Effective Error Angle for LF waveforms. Bmild search is used for WNB, and Iweak search for SG. Brown line is equal to 50% of y-axis.*

We report in Fig. 4.11 the median Effective Error Angle for WNB and SGQ9 LF waveforms. This figure shows that the value is not uniform over the sky, but it is affected by the source direction, and also by the tested waveforms. Results are similar for the other waveforms.

Obviously, the erA(0) value depends strongly on the SNR of the waveform: more SNR is larger, it is easier to identify the waveform, and conseguently all its characteristics. For small SNR, erA(0) assumes greater values, and increasing SNR it assumes an asymptotic value, which cannot be smaller than the angular resolution applied. In general, the asymptotic value is reached for SNR near $40 \div 60$ (Figs. 4.12 and 4.13). However, this SNR value is quite big respect to the one espected from a probable detection. An useful value is the Median erA(0) for SNR near to 20, as reported in the Tab. 4.4, where the values show that results are quite far from the expected.

In general higher frequency waveforms show better results than lower frequency ones. This is due to the fact that the likelihood peaks in the sky map are narrower for the high frequency cases. So, passing from one peak to the near one, the angular

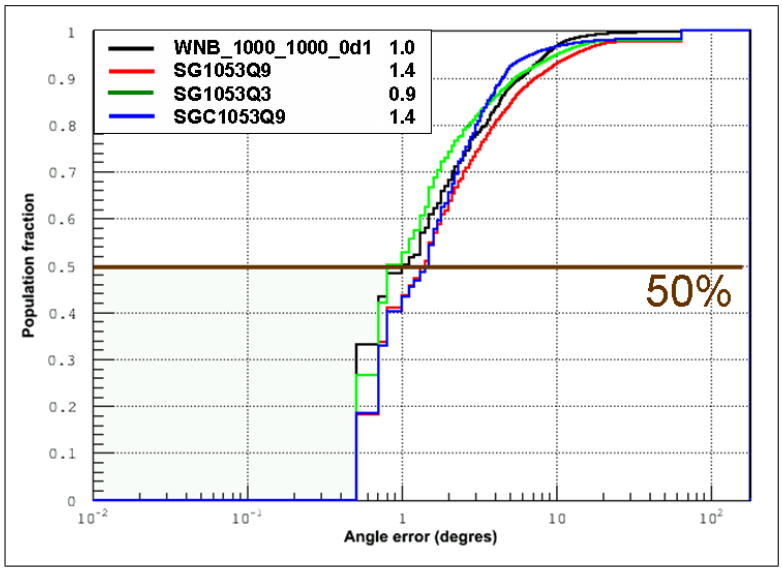

Figure 4.10: *Cumulative curves of Effective Error Angle for HF waveforms. Bmild search is used for WNB, and Iweak search for SG. Brown line is equal to 50% of y-axis.*

MDC Type		**Median ErA(0)** (°)			
Type	**Freq** (Hz)	**B mild**		**Best**	
		SNR=20	Asymptote	SNR=20	Asymptote
WNB	250	5.4	1.5	5.4	1.5
	1000	3.8	0.8	3.8	0.8
SGQ9	235	7.1	1.6	3.5	0.7
	1053	4.2	1.4	2.8	0.7
SGQ3	235	5.1	1.7	2.6	0.7
	1053	3.2	1.0	2.0	0.7
SGCQ9	235	7.2	3.3	2.0	0.5
	1053	3.6	1.4	1.2	0.6

Table 4.4: *Median ErA(0) from the fitting Eq. 4.6 for SNR=20 and Asymptotic value, B mild and Best search*

difference is lower for high frequency. This is why generally high frequency shows better direction reconstruction.

As espected, the Best choice shows better results than B mild. The asymptotic value for the best choice is about $0.6 \div 0.7$, and we cannot espect lower values than the angular resolution applied (0.5 degrees). The improvement of Best choice is visibile also for lower SNR, where we see a reduction of about two, especially for

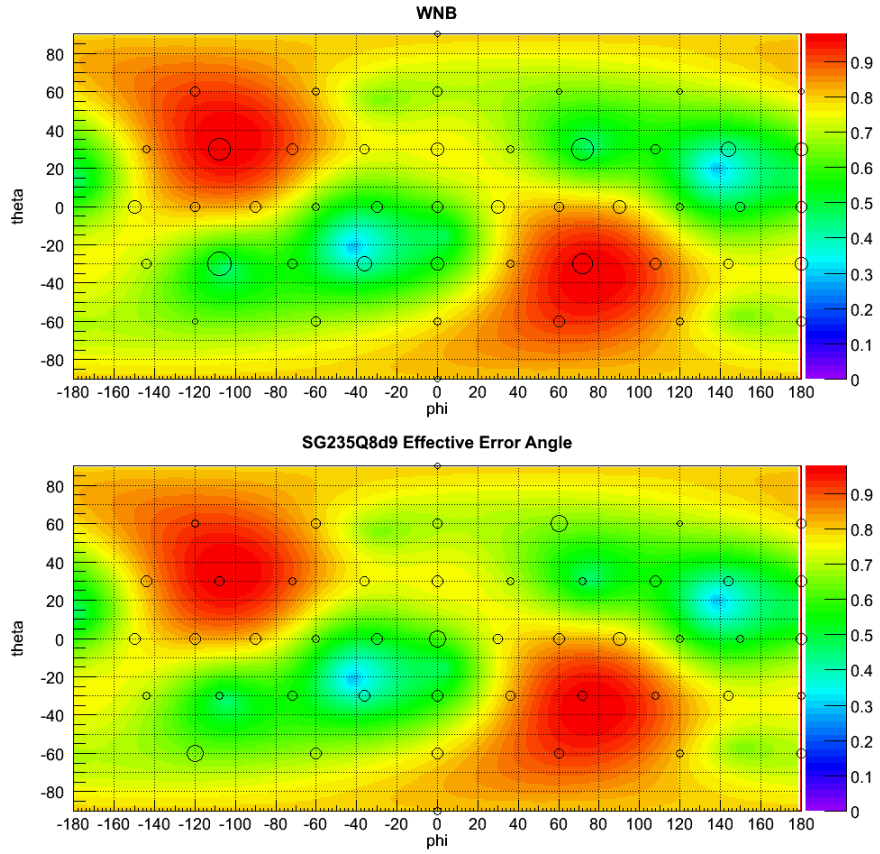

Figure 4.11: *Median Effective Error Angle for the 46 tested sky positions for WNB (top) and SGQ9 (bottom) LF waveforms: Bmild search is used for WNB, and Iweak search for SG.*

lower frequency waveform.

Moreover, results are not satisfying, the goal is to decrease the value of Effective Error Angle, i.e. to improve the reconstruction performances, such a way to give a more realible estimation of the reconstructed direction.

Error Region

The introduction of Error Region is an useful tricks to consider the reconstruction errors due to algorithm approximations and the network geometry: its definition is like an error area on reconstruction. However, the definition should satisfies two conditions:

- Well defined: the coverage of Error Region is corrected (Probability related

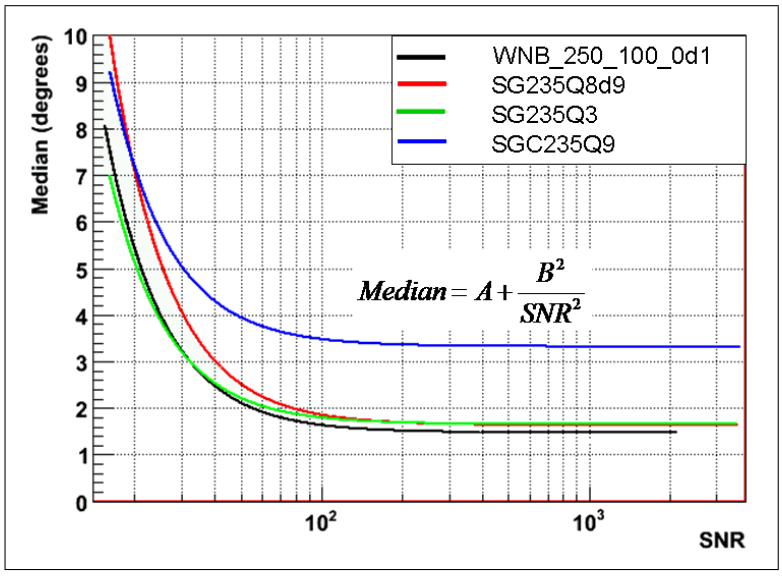

Figure 4.12: *Median erA(0) as a function of log(SNR) for Bmild search and different LF waveforms: WNB (black), SGQ9 (red), SGQ3 (green) and SGCQ9 (blue). The curves are fitting function following Eq. 4.6*

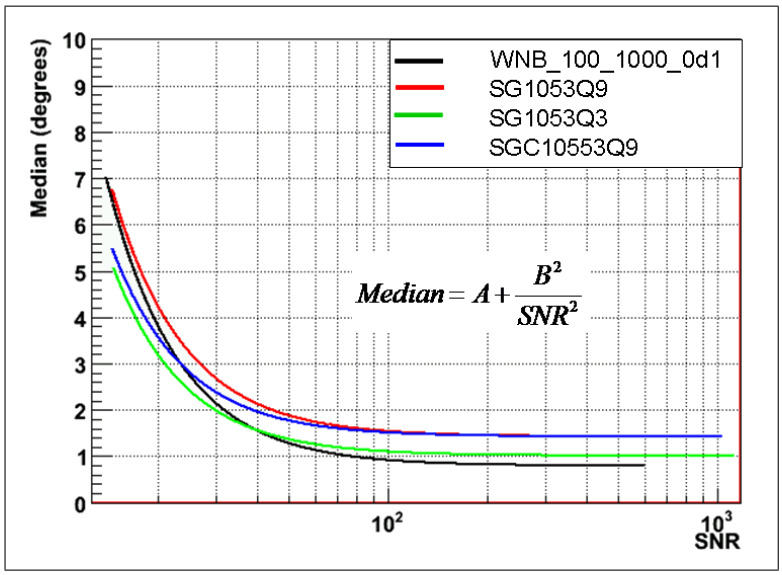

Figure 4.13: *Median erA(0) as a function of log(SNR) for Bmild search and different HF waveforms: WNB (black), SGQ9 (red), SGQ3 (green) and SGCQ9 (blue). The curves are fitting function following Eq. 4.6*

to Error Region X is equal to X%, and not more or less).

- The Error Region is enough small to be usable (a few degrees).

These two conditions can be verified only with simulation tests.

First of all we have to find a way to define Error Region for each reconstructed events. We have just esplained before how we can do it: we convert the Likelihood Sky Map to a Probability Sky Map introducing a conversion function (Sec. 4.2.5). We can change the values of K and pr to change the performances. For wat-5.0.9 tests the used values are:

- $K_0 = 10$
- $pr = 0.001$

First of all we can compare the Error Region behaviour with Effective Error Angle using cumulative curves (Figs. 4.14 and 4.15). These figures report the median value over all the directions of the Error Region (dashed curves) with respect to the Effective Error Angle (continuous curves). A good definiition of Error Region would make that the continuous and dashed curves cross at y-value equal to 0.5, so that the Error Region cover the 50% of the injections direction. Moreover, if the corresponding value on x-axis is small (a few degrees) this make the Error Region value usable. However, these conditions are quite satisfyed, because sometimes the Error Region is over estimated (the cross is over 0.5).

This fact could be accettable if the median behaviour is realized also for the single sky positions. In fact, the median value of Error Region reported in Fig. 4.16 does not depend on the source directions. This means that the Error Region, if well defined, could be usable, because, even if the reconstruction is affected by the network geometry, the Error Region does not care about it, so its performances are general over the sky.

To verify if the Error Region is right defined, we must consider its **coverage**. This quantity is defined as the fraction of reconstructed events for which the Error Region includes the injection position (i.e Error Region is bigger than Effective Error Angle). If the coverage is equal to 50% of reconstructed events, Error Region is well defined, otherwise it is under estimated (<50%) or over estimated (>50%). In Fig 4.17 we report the estimated coverage over the source directions. The figure shows that the quantity is strongly depending on the sky directions, which means that Error Region is not well defined over all the sky, but for some directions is under estimated, and for others is over estimated.

In conclusion, Error Region definition does not satisfy all the desired goals, and so it is not usable. This carries out to two possibilities: try to improve the

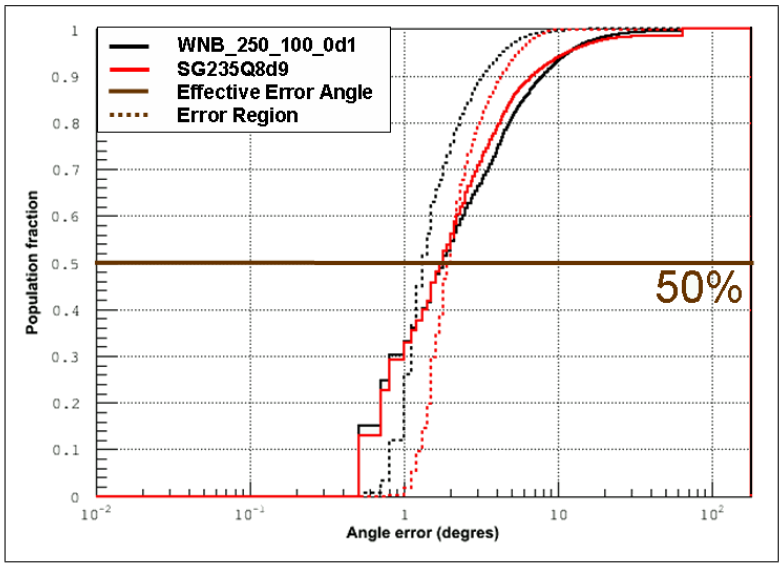

Figure 4.14: *Cumulative curves of Effective Error Angle (continuous) and Error Region dashed) for WNB and SGQ9 LF waveforms. Bmild search is used for WNB, and Iweak search for SG. Brown line is equal to 50% of y-axis.*

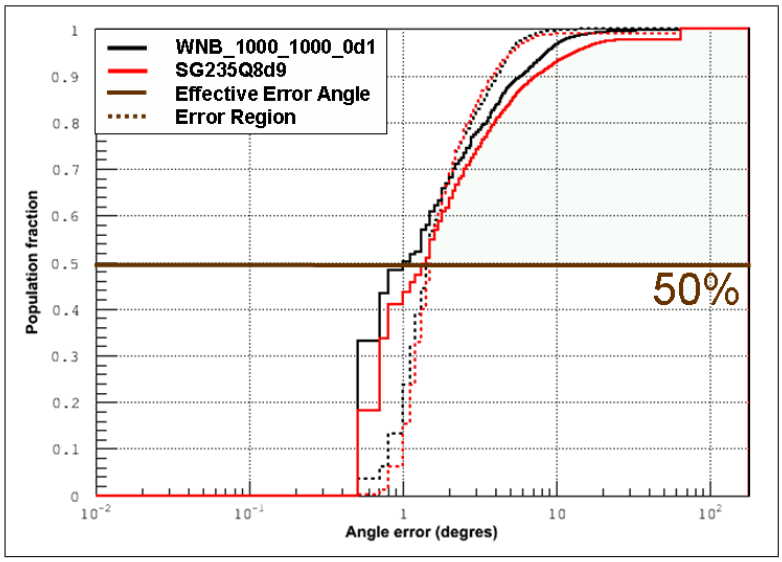

Figure 4.15: *Cumulative curves of Effective Error Angle (continuous) and Error Region dashed) for WNB and SGQ9 HF waveforms. Bmild search is used for WNB, and Iweak search for SG. Brown line is equal to 50% of y-axis.*

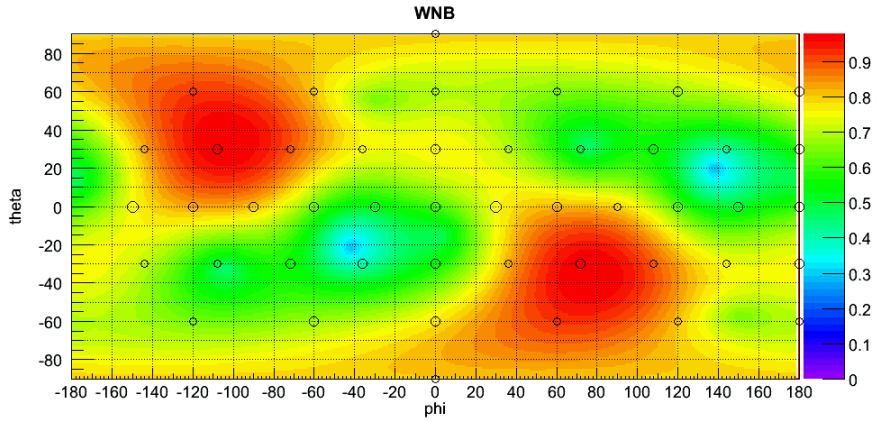

Figure 4.16: *Error Region for the 46 tested sky positions and WNB LF waveform, Bmild constrain.*

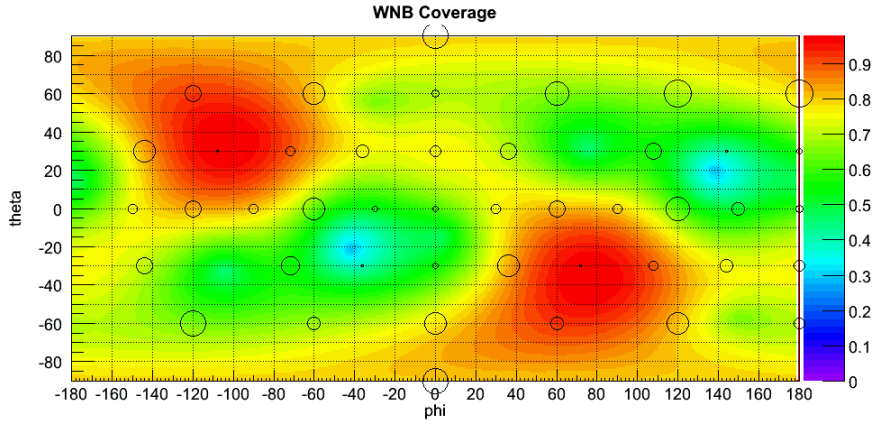

Figure 4.17: *Error Region coverage for the 46 tested sky positions and WNB LF waveform, Bmild constrain.*

Error Region definition to reach the desired goal, or leave behind this quantity and implement an alternative option to substitute it. A probable solution is to use Effective Error Angle as a "frequentist" Error Region defined from simulations. However, this holds on the limitations which we described above: this quantity is not general, because it is strongly dependent on the tested waveforms.

Chapter 5

LSC-Virgo Analysis

In this section we discuss the results of the first long term arch for GW bursts by the network of four detectors (V1, L1, H1, H2) belonging to Virgo and LSC collaborations. The version of coherent WaveBurst used to analyze these data is not as recent as the pipeline described in the previous section, which has been developed after this analysis; in particular, the features of reconstruction of signal parameters were not comparable and the investigation of the accidental backgound was much more limited. However, this older version was already well performing for what concerns the capability of identification of GW candidates and rejection of accidental events, being the most advanced pipeline available at that time with a computational efficiency suitable for the analysis of a long term observation over a broad frequency bandwidth.

The search is targeted to GW **bursts** in the frequency range of $2 \div 6$ kHz. We call this band **High Frequency** (**HF**) in opposition with the Low Frequency (LF), which considers frequencies lower than 2 kHz (Fig. 5.1). The CWB version used for this work is the WAT 4.7.3 [101], adapted to run over a larger frequency band than previously used. This has been the first coherent search in the HF band and its responsibility was assigned to our Virgo group.

We describe in the following the data stream analysed (S5/VSR1), the MDC tested for the efficiency results and upper limits, the tuning procedure of the post-production thresholds and finally the results on zero lags.

5.1 Data set

The high sensitivity required for the GW detection implies that data are significantly affected by noise (Chapter 2.1). Interferometers are continuously checked and commissioning work is regularly planned. When a particular test is performed, or the instrument is off to permit an hardware modification, the data stream is

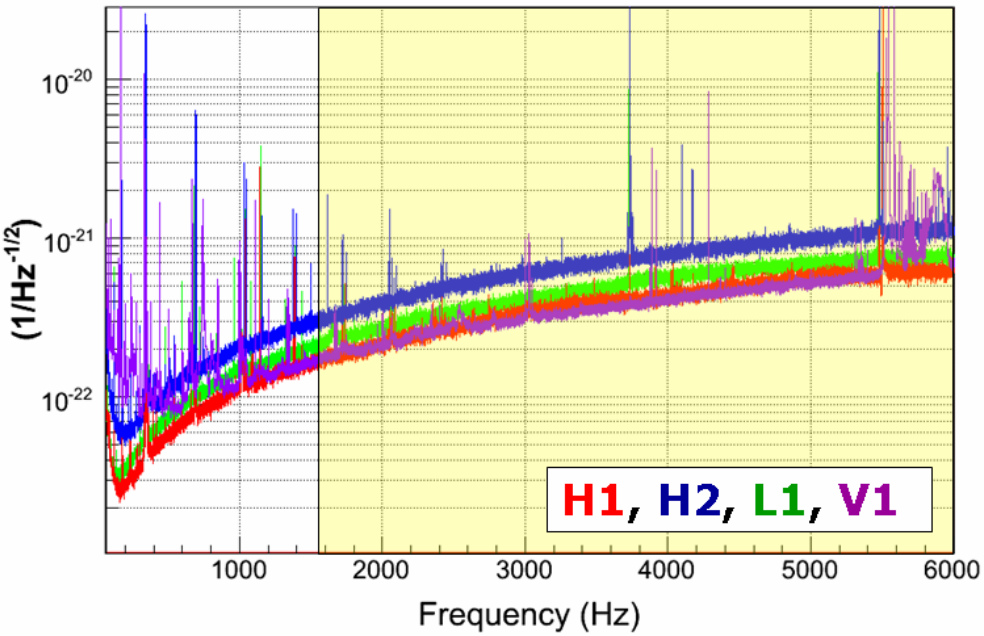

Figure 5.1: *Sensitivity curves for H1 (red), H2 (blue), L1 (green), V1(purple) during JW1, a week joint period of data acquisition between Virgo and LSC. At 6 kHz the LSC data show some sort of a "cut-off" probably due to the calibration.*

not usable for the detection of GWs. We use the word **Science Mode** to identify data which can be used for GW detection.

A science run is a definite time period where the detector is in science mode. Start and stop run times are set to meet the scientific motivations decided for that run. This work uses the data stream belonging to the second year of the fifth LSC run (**S5Y2**) and the whole first Virgo run (**VSR1**) according to the agreement between LSC and Virgo collaboration. We report in the Tab. 5.1 the starting and ending times of these scientific runs.

Run	Start		Stop	
	GPS	UTC	GPS	UTC
S5Y2	847555214	2007-11-14 16:00:00	875289614	2007-10-01 16:00:00
VSR1	863558950	2007-05-18 21:00:00	875253398	2007-10-01 05:56:24

Table 5.1: *Start and stop in GPS and UTC times of the S5Y2 and VSR1 runs.*

An subperiod of the S5Y2/VSR1 has been selected to perform pilot tests, the so

called **JW1** (Joint Week 1, Tab. 5.2), about five days of coincidence time between the four detectors (V1, H1, H2 and L1). In fact, this small period is representative of the whole S5Y2/VSR1 for what concerns the detection efficiency. Therefore, it can be used to compare the sensitivity of different settings of CWB parameters as well as the robustness of CWB against calibration systematic errors with a limited computational load perform mis-calibration tests, or to tune CWB parameters for the analysis (This allows to decrease the computational load). Unfortunately, JW1 background is not representative of the whole S5Y2/VSR1 period, so we cannot perform background tests on JW1 period and extend result considerations on the whole S5Y2/VSR1, as for efficiency tests.

Run	Start		Stop	
	GPS	UTC	GPS	UTC
JW1	870052940	2007-08-02 01:22:06	870623383	2007-08-08 15:49:08

Table 5.2: *Start and stop in GPS and UTC times of the JW1 period.*

The data stream calibration to reconstruct h(t) is a remanding task, and it is very important for the correct use of a coherent data analysis to identify and estimate GW parameters. However, the calibration is not static, but revisions and adjustements produces new versions when possible. For this work we used the most recent calibration version available at the time of the analysis: the **RDS_C03_L2** for LSC, the so called **V3** [102], and the **Hrec_V2**, also known as **V2** [103] for Virgo. More details on the calibration are on the Sec. 5.1.2.

5.1.1 Data Quality

When a detector is in data taking mode, not all the data stream are considered of good enough quality to be used for searching GW. The time list of data stream that can be used for the analysis is provided by the **Detector Characterization and Data Quality Groups**, according to the behaviour of the different detector channels. The various time lists referring to useful/unuseful data stream can be found using the web interface provided by the group [104].

Lists can be collected in four main classes: *Science Mode*, *Injection Mode*, *Locked Mode* and *Data Quality* [105]. Each list is organised in *segments*, i.e. time periods of variable length which define the GPS times in which detector data stream are in good or bad quality. Start and stop times for each segment are defined with 1 second resolution.

The informations stored for each list segments are:

- run tag (VSR1,...)

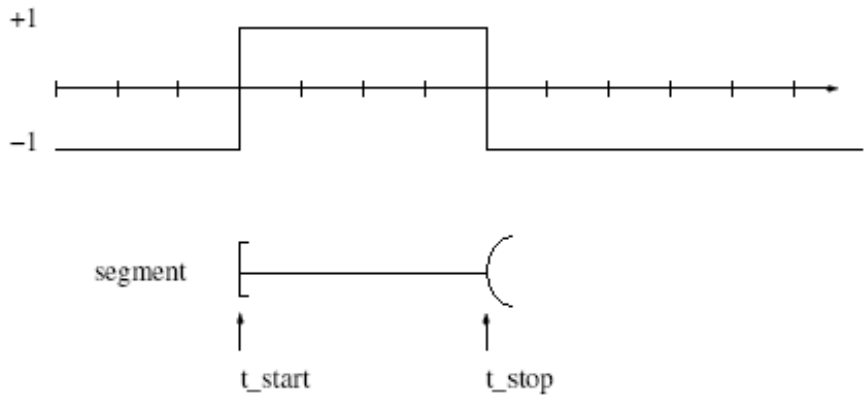

Figure 5.2: *Definition of the segment's boundaries. Starts and stops are respectively included and excluded from the segment.*

- a progressive number (1, 2, 3)
- GPS start time
- GPS stop time
- state (-1, 0, 1)

where the state value is set according to Tab. 5.3

List type	+1	0	-1
Science Mode	science mode	not science mode	not defined
Data Quality	bad quality	good quality	not defined

Table 5.3: *State definition for Science Mode and Data Quality segments. Not defined means that data have not yet been tested.*

The output can be reported also in an ASCII file, in this case the list contains the following values:

- segment number (1, 2, 3)
- GPS start time
- GPS stop time
- duration (optional)

We report now some informations on the different segment classes, focusing on Virgo detector.

Science Mode

Science Mode segments are the baseline lists for all data analysis. Segments belonging to this group are defined by the value of the Qc_Moni_ScienceMode channel, as recorded in the trend data frame files. This channel is derived online from the raw data, to be reliable and faster to read.

Definition of science mode is static (so there is no version number).

If the case of missing frame data or missing Qc_Moni_ScienceMode data, the science mode status is defined as the last available status before interruption, except if the detector leaves science mode during the period of missing data.

In some cases, missing data segment could be long enough that the detector could leave and return to science mode during a single instance of missing data. In these cases the science status is manually double checked during the periods of missing data. If it is found that data are incorrectly marked as science mode, this is handled by a specific data quality flag.

Injection Mode

This group identify time periods when simulated GW signals are physically injected into the detector. A flag (Alp_Cali_EVentInj_STATUS) is produced and stored in the raw data files. The missing data policy is similar to the one defined for the science segment. The list of injections details (time, SNR, ...) are mantained in the database indipendently of the segment list.

Locked Mode

This list corresponds to the time period when the detector is locked (see Sec. 2.1). These segments could be useful for running online analysis in order to support commisioning activities. They could be interesting also to follow-up interesting candidates identified by other detectors. The channel used is the Alp_Main_LOCK_STEP_STATUS. When this channel is set equal to 12, the flag in Virgo DataBase is set equal to 1, otherwise it set 0. If the channel is not available, the flag is set equal to -1. It is possible that for server problems the interfometer is not locked while Alp_Main_LOCK_STEP_STATUS=12. In this case a specific quality flag is produced to tag these time periods.

Data quality

The purpose of these segments is to identify periods of poor data quality or known problems that affect science mode segments. Some examples of data quality flags are:

- incorrect science mode
- unmarked injections
- missing data
- bad h reconstruction
- airplane or helicopter overflight
- electrical power glitch

Data quality segments are used in conjunction with science mode segments to restrict the list of usable segments. These flags can be related to different thresholds (WIND_OVER_40KPH, WIND_OVER_50KPH) or different frequency bands (IBZ_BRMS_60_300HZ, IBZ_BRMS_600_700HZ), in this case they are implemented as separate flags. The newest version of a dataquality flag supercedes all previous versions and the database returns the most recent version unless a user explicitly requests an earlier version.

Data Quality flags are collected in several categories depending of the type of usage and the characteristics (mainly efficiency and dead time):

- **Category I**: flag should be applied on data segments before event processing since we expect abnormal data characteristics that could jeopardize the performances of the pipeline in adjacent periods.
- **Category II**: flag can be applied on the resulting triggers during the analysis or during the production stage. This category corresponds to well understood misfunctioning of the detector. The deadtime remains small (few %).
- **Category III**: flag should be applied to veto triggers after the analysis. It flags disturbances in environmental and instrumental auxiliary channels which are statistically correlated to excess noise and triggers on the gravitational wave channel, though their casual relation is not proved. Dead time may be large and the interest of the DQ depends on the analysis. This category flag is used to select a subset of triggers with greater overall confidences as GW candidates; however, any single trigger vetoed by a cat 3 cannot be considered a fake signal, given the statistical nature of the relationship with

disturbances. In fact, the use percentage, i.e. the fraction of Category III vetoes which succefully kill triggers in the GW channel is low. This category could help in assessing an a posteriori confidence on plausibility for a GW candidate.

We can include in this Category also the **event by event vetoes** [106] which are applied in the post-processing phase of the analysis. However, veto lists are different from DQ, in fact the start and stop times are not integer numbers, but they reach the ms resolution. This because veto lists are constructed in different way and with the aim the select precisely defined time periods which are particularly noising.

The same flag can be put in different categories for different analysis, however there is a "recommended" category for each flag and standard analyses.

For this work we considered the **Virgo VH2** [107] and the **LIGO DQ 2.5** [108]. Category I segments have been used to define the job segments, whereas Category II have been used to define time periods during production analysis.

Data Quality flags are set separately for each considered detector, their intersection define the livetime of network configurations made by different sets of detectors. In the Tab. 5.4 we report the *exclusive*[1] live times for all the possible network configurations considering the four detectors after applying the LSC and Virgo DQ Category II.

We have decided to analyses network configurations with sufficient duty time (in red in the Tab. 5.4), including the H1L1V1 network, for the possibility for this configuration to increase statistic adding a fixed shift on the second detector (Section 5.3.2).

5.1.2 Calibration Issues

The reconstructed h(t) data stream is affected by a frequency dependence calibration uncertainty for the amplitude and phase (or time offset). The amplitude uncertainty is reported as estimated deviation plus its one sigma estimate. In fact, different calibration measurements show that the mis-calibration in amplitude of each detector apparently fluctuates from time to time in a random way. It is not known how much of these fluctuations are due to the actuators which force the mirrors to inject reference signals during the calibration measurements or instead to a true change in calibration of the GW sensing of the interferometer.

[1]The H1H2L1V1 concidence period is obviously inside the H1H2L1 one. The H1H2L1 esclusive time is defined as the entire H1H2L1 coincidence time excluding the H1H2L1V1 time. Same definition is applied for the other networks.

Number of detectors	Network	Live time [days]
4	H1 H2 L1 V1	68.9
3	H1 H2 L1	124.5
	H1 H2 V1	15.8
	H1 L1 V1	4.5
	H2 L1 V1	1.3
2	H1 H2	35.4
	H1 L1	7.2
	H2 L1	3.8
	H1 V1	1.8
	H2 V1	0.8
	L1 V1	6.3

Table 5.4: *Exclusive live times for all network configurations during S5Y2/VSR1 period. In red the network that have been analysed for this work.*

These fluctutations of amplitude calibration should than be considered because they may affect the estimation of the efficiency and upper limits. To consider this fact simulation tests are made injecting "ad-hoc" simulated signals. In practice we mimic the effect of calibration systematics on the detection eficiency of the network by injecting sotware GW signals with uncalibrated amplitude and phase charcteristics in different detectors (Sec. 5.4). In this work we considered two effects:

- the use of not final calibration versions. In fact, when we started the analysis, the LSC V4 and Virgo V2 calibrations were not ready. Luckily, the Virgo V2 was ready in time to re-run the analysis, but for LIGO detectors we had to stick to the V3 calibration and then take into account the known systematics between V3 and V4 calibration [109] as a sistematic uncertainty on the V3.
- the rated uncertainties of the final calibration versions. Virgo reports a simple upper limit on the expected amplitude and phase/time calibration systematics. LIGO reports the rms taken from the calibration measurements.

The systematic errors expected for Virgo V2 [103] are:

- amplitude: 6%
- time: 6μs above 1.9 kHz

These are meant as upper limits, but to be conservative we considered them as standard deviations.

Detector	**Systematic deviation** ((V3-V4)/V4 %)	**Rms (%)**		
		0-2 kHz	2-4 kHz	4-6 kHz
L1	[2.5, 5]	13.4	13.3	13.5
H1	[-8, -5]	10.2	15.4	24.4
H2	[-11, -9]	10.3	11.3	16.6

Table 5.5: *Systematic difference versus frequency of V4 calibration (final) and previous V3 calibration and rms fluctutation of V4 calibration for the LSC detectors.*

The amplitude and phase errors for the LSC detectors are shown in Fig. 5.3. The blue line in the amplitude plot shows a problem on the H2 detector, but we are not interested in it because it is outside the frequency band of interest ($2 \div 6$ kHz). The red line in the phase plot indicates a problem on H1 detector: a missing filter in the h(t) production causes a very large phase systematics (ten of degrees) which is fortunately limited in frequency (some Hz). To resolve this problem we apply a post-production cut on all the events occurring with a reconstructed frequency of a few Hz around the H1 resonance. CWB estimates this resonance at 5482.5 Hz, so we cut events with reconstructed frequency within 5482.5 ± 2.5 Hz.

We report in Tab 5.5 the calibration differences between V3 and V4 versions, following the Calibration team instructions [109], which report two types of contributions: the systematics and the statistical rms. The former indicates a systematic difference between V3 and V4 calibrations which depends on the frequency. The latter is taken as one sigma estimate, over three frequency bands:: lower than 2 kHz, from 2 to 4 kHz, and higher than 4 kHz.

5.2 Mock Data Challenge

The performances of a generic algorithm of GW search are estimated from MonteCarlo simulations. These tests consist of injecting simulated waveforms in the data stream and estimating the efficiency reconstruction of these injected signals. These are *software* injections, because they are made after the data taking. Also *hardware* injections are performed during the data taking, but for this work we do not consider them, in fact there is a proper DQ flag that avoids time periods containing hardware injections. Besides, the hardware injections have been mostly performed without mimicking an actual GW signal, and therefore are not useful to test a coherent data analysis pipeline.

To perform software injections in the right way, the GW community has established a common method to exchange simulated waveform data stream for the several detectors: the **Mock Data Challenge** (**MDC**). In fact, given a source di-

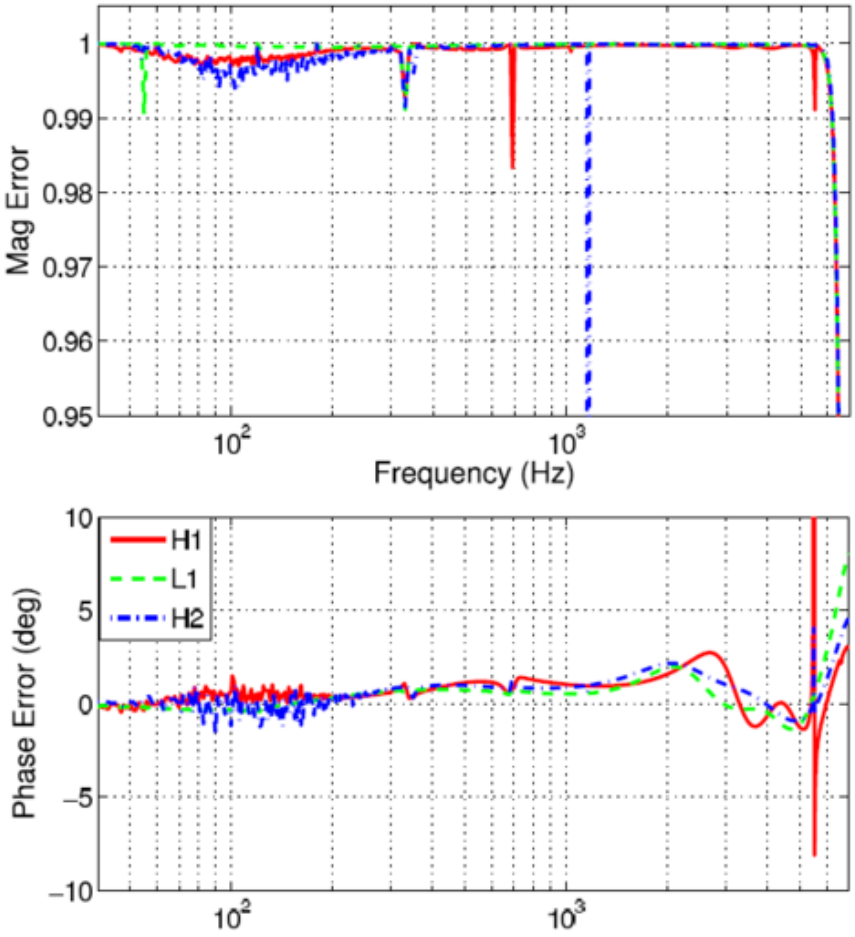

Figure 5.3: *Amplitude and phase systematics as a function of frequency for the three LSC detectors. Blue line in amplitude shows a problem on the H2 detector. Phase error shown in red at 5.5 kHz for H1 indicates a problem with the filters used in the h(t) production.*

rection in the sky, the wavefron will hit the detectors at different times. Morevoer, each detector senses a different combination of + and × polarizations, due to the different orientations of the detector's arms (Eq. 2.3). The MDC data stream contains a set of gravitational waveforms to be added in each detector data stream listing the following parameters:

- the source direction
- time arrival at each detector, and at the earth center

- h_+ and $h_\times$ polarization waveforms at earth
- F_+ and $F_\times$ directional sensitivity components for each detector
- h_{rss} values at the earth

MDC are produced by a software developed in the LIGO group: the **Burst-MDC** package [110].

Obviously we cannot test the pipeline over all possible GW Burst Waveforms nor we can provide any "complete" set of waveforms.

We have produced two sets, which we have identified with the label SGQ9 and MDC2 and we explained in details in the following sections.

5.2.1 Sine Gaussian MDC set

This set is composed of Sine Gaussian waveforms. The general formula of a Sine Gaussian is:

$$h = e^{-(t/\tau)^2} Sin(2\pi f t) \tag{5.1}$$

where the decay time depends on central frequency f: $\tau = \frac{Q}{\sqrt{2}\pi f}$. So usually the set is identified with the value of Q chosen.

SineGaussians have been chosen because thay are compact signals in the time-frequency plane. Moreover, by varying the central frequency and Q value, they can probe systematically a reasonable large burst parameter space.

For our test we use the **HSG7_S5_R3** version [111], with the following characteristics:

- 13 central frequencies, approximately equally spaced in logarithmic frequency in the HF band (Tab. 5.6 and Fig. 5.4). The choice of frequency was fixed in a previous analysis.
- Q=9, if $f \approx 1415 \div 6000$ Hz, $\tau \approx 0.34 \div 1.43$ ms.
- Reference strength: hrss=$2.5 \cdot 10^{-21}$ $1/\sqrt{Hz}$.
- Injections every 100 s, injection time randomly chosen within $\pm$ 10 s of center.
- Random selection of the 13 waveforms at each injection time, so that the mean rate of each waveform is 1/1300 s.
- Input waves are *linearly polarized signals*.
- Simulation randomly selects sky directions from a uniform distribution in the sky and polarization angle from a uniform distribution in [0, ψ], therefore getting a random combination of h_+ and $h_\times$ polarizations at earth.

We distinguish each waveform indicating the central frequency and the Q. So, a generic SineGaussian is named SGXXXXQY, where X is the central frequency, and Y the Q.

SGQ9				
SG1415Q9	SG1615Q9	SG1797Q9	SG2000Q9	SG2226Q9
SG2477Q9	SG2756Q9	SG3067Q9	SG3413Q9	SG3799Q9
SG4225Q9	SG5000Q9	SG6000Q9		

Table 5.6: *List of waveforms belonging to SGQ9 set.*

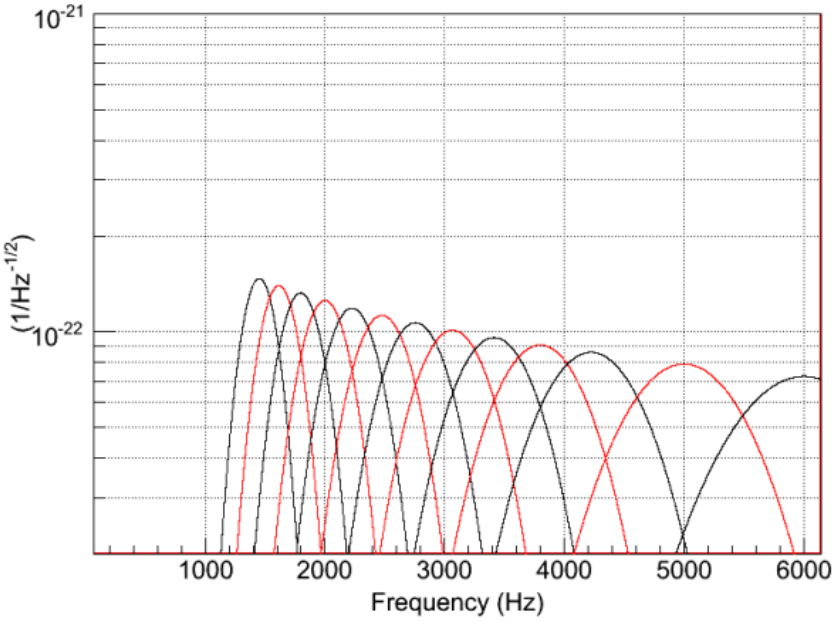

Figure 5.4: *Fourier Transform of SGQ9 set. Frequencies are spaced according to the bandwidths of the waveforms.*

This set has been used to tune the post-production parameters of the CWB for the several network configurations (Chapter 5.3.4).

5.2.2 Second MDC set

It is preferable to test the pipeline also on other waveforms with different time-frequency characteristics, to mimic a larger fraction of possible astrophysical pro-

cesses. For this purpose we have prepared a second MDC set, which contains the following waveform:

- Sine Gaussian (**SG**) wih shorter and longer duration;
- White Noise Bursts (**WNB**);
- Ring Downs (**RD**);
- Numerical Supernova Waveforms (**SN**).

We planned to use a unique set including these waveforms to reduce the computational load, mantaining a sufficient statistic to obtain reliable efficiency curves. However, some technical problems with BurstMDC force us to divide the waveforms in two separate sets (**HBRST1** and **HBRST2**), so we decide to halve the injection rate of signals (1 every 200 s instead of 100 s) to mantain the planned computational load. Other characteristics are the same for the SGQ9 set. In the Tab. 5.7 there is the list of waveforms for the two sets.

HBRST1					
RD	RDL_200_1590	RDL_200_2090	RDL_200_2590		
	RDC_200_1590	RDC_200_2090	RDC_200_2590		
	RDL_Q9_2000	RDL_Q9_3067	RDL_Q9_5000		
	RDC_Q9_2000	RDC_Q9_3067	RDC_Q9_5000		
WNB	WNB_2000_1000_0d01	WNB_3500_1000_0d01	WNB_5000_1000_0d01		
	WNB_2000_100_0d1	WNB_2500_100_0d1	WNB_5000_100_0d1		
HBRST2					
SG	SG1615Q3	SG2000Q3	SG2477Q3	SG3067Q3	SG5000Q3
	SG1615Q100	SG2000Q100	SG2477Q100	SG3067Q100	SG5000Q100
SN	D1	D4			

Table 5.7: *List waveforms belonging to MDC2 set.*

Sine Gaussians

The name convention for the SineGaussian are the same as the SGQ9 set. In this set we use waveforms with different Q: 3 and 100. In the range central frequency $1415 \div 6000$ we obtain decay times for Q=3 (100) between $0.11 \div 0.47$ ($3.75 \div 15.90$) ms. These waveforms are used to test the pipeline on different duration and bandwidth scales with respect to SGQ9.

White Noise Bursts

These waveforms are Bursts built by bandpassing in a selected bandwidth a white noise of selected duration with independent polarization amplitudes. These are considered as the limiting case signals, since they have a close resemblance to the noise of each detector and stand up in the network data analysis just because they partially correlate the detector responses. The name convention is: WNB_startfreq_bandwidth_duration. Each waveform type is made by a pair of random h_+ and indipendent random $h_\times$. This is accomplished bt sampling randomly the h_+ and $h_\times$ waveforms from a set of 30 scalar waveforms, therefore obtaining 30x29 independent $h_{+/\times}$ couples.

Ring Down

Ring Down waveform are expected from processes like f-modes of NS, or during the Ring-Down of newly formed BH/NS during the post merger phase after coalesce (Chapter 1.2.1), or produced in SN star-quakes related to Soft Gamma Repeaters (SGR). They are characterized by frequency, τ or Q (similarly to SG) and polarization. The name convention is: RD*_tau_frequency or RD*_Q_frequency, where the * indicates the polarization: "L" stands for linear polarization ($h_\times = 0$) and "C" for circular polarization.

The expression for Ring Down waveform with circular polarization is:

$$\begin{pmatrix} h_+ \\ h_\times \end{pmatrix} = \frac{h_{rss}}{\pi f \tau}\sqrt{\frac{1+4\pi^2 f^2\tau^2}{\tau(1+e^{-1/(2f\tau)})}}e^{-t/\tau}\begin{pmatrix} Cos2\psi & -Sin2\psi \\ Sin2\psi & Cos2\psi \end{pmatrix}\begin{pmatrix} \frac{1+cos^2\iota}{2}\Theta(t-\frac{1}{4f})Cos(2\pi ft) \\ Cos(\iota)\Theta(t)Sin(2\pi ft) \end{pmatrix} \tag{5.2}$$

where we have defined:

$$h_{rss}^2 = \int_0^\infty (|h_+(t)|^2 + |h_\times(t)|^2)dt \tag{5.3}$$

the angle ψ is an arbitrary GW polarization uniformly distribuited in [0, π] and the angle ι an arbitrary inclination of the system wich originates the burst with respect to the line of sight. If we assume a random orientation in space of the axis of symmetry of the emitting system the cos(ι) should be uniformly distribuited in [-1, 1]. The Θ function is introduced to avoid a discontinuity at the beginning of the waveform which would result into a infinite energy. Linearly polarized RD consider only the h_+ component in the expression above

SN waveforms

These waveforms are predicted from processes of SN core collapse. We have considered two waveforms among the great variety in literature, which we indicate

with the name **D1** and **D4** [112]. The first is generated by the collapse of a 1.26 solar mass star that is almost spherical at the time of the transition into a BH, while the latter is generated by the collapse of a 1.86 solar mass star that is maximally deformed at the time of the transition into a BH. Only the h_+ component is considered, because the $h_\times$ is smaller by more than an order of magnitude (Fig. 5.5).

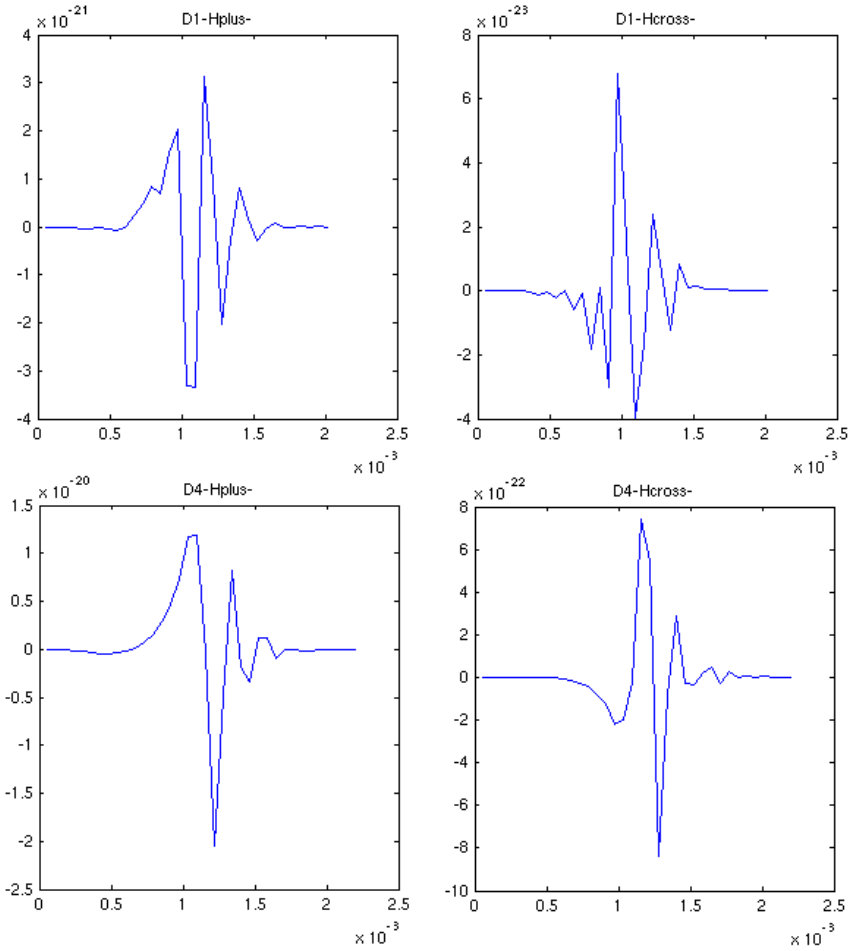

Figure 5.5: *D1 (up) and D4 (down) waveforms in time domain, plus (left) and cross (right) components. The possible sources are at a distance of 10 kpc.*

All the waveforms are generated with a reference amplitude $h_{rss} = 2.5 \cdot 10^{-21}$ $\mathrm{Hz}^{-1/2}$ except for the SN. In this case the signal amplitude is set by the numerical simulations of the source, so that changing the injected h_{rss} at earth is equivalent

to scaling the source distance. The reference h_{rss} generated for SN are $h_{rss} = 5.7 \cdot 10^{-23}$ for D1 and $h_{rss} = 2.5 \cdot 10^{-22}\ 1/\sqrt{Hz}$ for D4, and they refer to a distance from the source at 10 kpc.

5.3 Analysis procedure

The goal of the analysis is to verify if in the S5Y2/VSR1 data stream there are candidate GW events, or to estimate upper limits on the rate of gravitational waves versus their strenght if there are no candidate events.

This goal is reached using a *blind* search: the decision rule of a candidate event is established a priori, and then applyed on the actual results of the data analysis. This decision rule is defined as a list of post-production cuts (Chapter 3.3), chosen according to the noise behaviour and the involved detectors.

Increasing or decreasing post-production thresholds affects the balance between false alarms and detection efficiency, the optimal choice should give a good compromise between an high detection efficiency and a low false alarm rate (section 5.3.4). Usually, false alarm are investigated by performing the same analysis on data streams shifted in time with respect to one another (section 5.3.2), so to cancel correlations due to any physical excitation of the network. Detection efficiency, instead, is computed using MonteCarlo signal injections in the data stream (section 5.3.3).

5.3.1 Down sampling

The initial sample rate of the data stream is equal to 16384 Hz. This would produce an high computational load. For the LF search this problem was solved selecting a sub-band of interest using wavelet transform. In fact we know that a chosen wavelet decomposition level carries the complete information about the signal. We can use this property to bandpass the data in the wavelet domain and inverting back in the time domain. However, this procedure is easily implementable when: the bandwidth is a fraction $1/2^n$ of the starting sample rate; the starting frequency is 0 or a multiple of the bandwidth. For instance, in the LF search, the frequency band of interest is 64-2000Hz, we can use wavelet transform to select the band 0-2048 Hz.[2]

The frequency band of interest for this work is over 2 kHz. LSC data have a "cut-off" at 6kHz, so we have to use this as a superior limit. In these conditions, it

[2]We select the 2^{nd} level, which has a frequency resolution equal to 2048. We have four groups of coefficients related to respective band: 0-2048, 2048-4096, 4096-6144, 6144-8192. We select only the coefficients related to the first one and we apply inverse transform.

is not easy to use wavelet transform to choose the frequency band. So, we perform a Down Sampling procedure in C++ language. The procedure can:

- Trasnform data in the frequency domain from a starting sampling rate to a lower sampling rate
- Make a shift in frequency from a chosen lower value to 0
- Select a frequency band

The procedure accepts four parameters:

- Final sampling rate
- Shift in frequency
- Start frequency of the band (after the shift, usually this is null)
- End frequency of the band (after the shift)

The procedure is simple:

- After some checks on the input values, we apply a Discrete Fourier Trasnform.
- The coefficients of the Fourier Trasnform are saved with an array of complex numbers
- The array coefficients are moved of the shift frequency
- Coefficients outside the Start-End Frequencies are put to 0
- Inverse Fourier Transform is applied to recover bandlimited data in time domain

In a first phase the procedure was performed to select the band between 1280 and 6000 Hz with a sampling rate of 10240 Hz (**5kHz-DC**). We selected this sampling rate to get multiples of the frequency resolutions at all the decomposition levels used.[3] The superimposition with the LF search is chosen to have a cross-check on the efficiency.

Obviously, any actual down sampling procedure have effects on data. Checking these effects we realise that the systematics introduced in the power spectra are negligible, but we found some problems on phase effects.

We report in Fig. 5.6 two examples of the amplitude effects on strain spectra (left) and a SineGaussian (right) in the frequency domain. Blue curve is the

[3]Because we use decomposition levels from 4 to 9, the frequency resolutions from a sampling rate of 10240 are from 10 to 300 Hz.

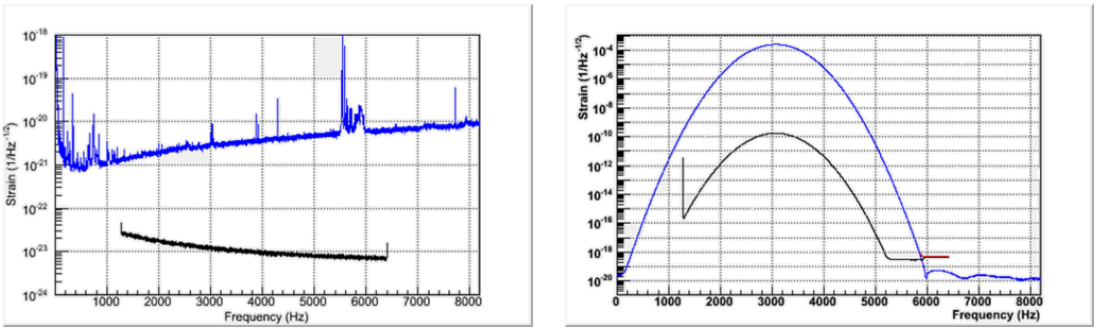

Figure 5.6: *Effect of 5kHz-DC on amplitude for Virgo spectra (left) and SG3067Q9 (right). Blue curve is the amplitude of original data (16384 Hz) applying a Blackman-Harris window, the black curve represents the differences between FFT converted to the sample rate of 10240 Hz and the original one (16384 Hz).*

original amplitude applying a Blackman-Harris window, the black curve represents the differences between FFT converted to the sample rate of 10240 Hz and the original one. Effects are small by orders of magnitude.

Unfortunately, the Down Conversion has a great distortion on the carrier part, whose frequency is shifted by the value of the start frequency. For instance, a SineGaussian with initial central frequency equal to 2000 Hz is down converted to a central frequency of 720 Hz (2000-1280). This is not a big problem, because the signal envelope is preserved, the waveform mantains the same duration and the arrival time does not change in a significant way. The problem is that the signal shape is modified depending on the phase of arrival of its carrier with respect to the reference frequency of the down-sampling (1280 Hz). If the SineGaussian phase is incoming $\pi/2$ later with respect to the 1280 Hz reference, the down converted signal becomes a CosineGaussian (Fig. 5.7).

We cannot neglect this effect because CWB considers the signal shape to make coeherent considerations and reconstruct the waveform. In practice, the reconstruction of the signal characteristics, e.g. the source direction and signal waveforms, turns out to be severely jeopardized and the resulting sensitivity gets worse, typically requiring signal amplitudes greater by 10% to get the same detection efficiency with respect to the analysis without downsampling. To avoid this problem we had to eliminate the frequency shift, and we chose a final sample rate of 12800 Hz, selecting all the band between 0-6400 Hz. The choice of frequency band of 1280-6000 Hz is made using CWB selections in the wavelet domain. The sample rate of 12800 Hz has been chosen making the same considerations about frequency resolutions of the decomposition levels. Using this sample rate, and decomposition levels from 4 to 9, we obtain a frequency resolution from 12.5 to 400 Hz (and a

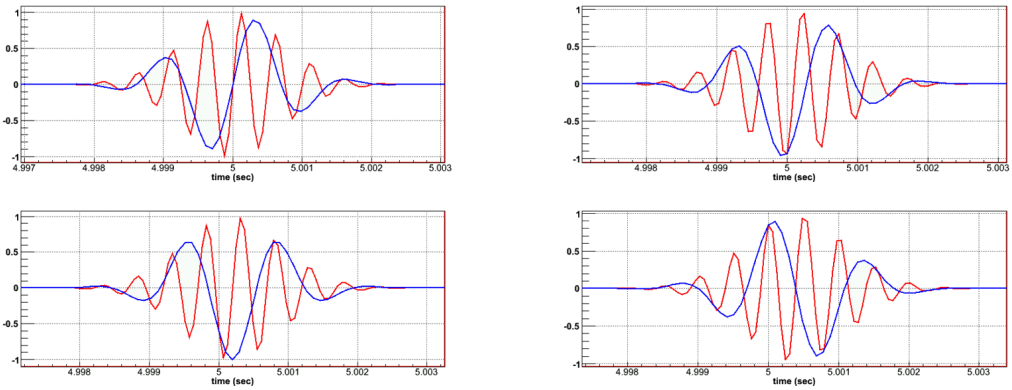

Figure 5.7: *Effect of 5kHz-DC on SG2000Q9. Red is the original signal, blue the down converted one. Signal is arriving at different phase after the buffer center:* $\pi/4$, $\pi/2$, $3\pi/4$, π *(from left to righ, top to bottom). The shape is modified for different phases.*

time resolution from 40 to 1.25 ms).

5.3.2 Background study

We have to establish a procedure to characterize statistics of the noise outliers estimated by the pipeline when analysing the data. Usually, detectors data are shifted of a time offset, so detected coincidence events cannot be related to a GW signal. If ΔT is the maximum light time delay between detectors between L1 and V1, we shift, for instance, one detector data stream of more than ΔT, so coherent events estimated by the pipeline cannot be GWs. Repeating many times this procedure with different time shifts (called time lags) we can characterize the distribution of the events occuring coherently in the network just by chance, in particular the accidental distribution of events occurring in the *non* shifted data (*zero lag* or *on-source data*).

We have esplained in the previous chapter (3.2.4) how the algorithm performs time shifts on data stream. Time shifts are applied only to the first detector of the network and the time shift step must be a multiple of the time resolution of the TF map. Starting from a sampling rate equal to 12800 Hz, and using the maximum decomposition level equal to 9, the worse time resolution is equal to 40ms. Maximum light time delay between detectors is about 27 ms involving the couple L1-V1. We selected a time shift step equal to 1.56 s, which is a multiple of 40 ms and it is sufficiently higher than the duration of events in the detectors, so to avoid correlations in adjacent time lags due to multiple superposition of the same events. The number of shifts used is connected to the maximum time shift

we can apply to data, i.e. 300 s, which is the minimum length of a job segments. In fact, as discussed in Chapter 3.2.4, the reference detector is shifted circularly within the segment and a shift equal to segment duration would just be equivalent to the zero lag. According to these considerations, maximum lag number used is 192 ($192 \cdot 1.56 < 300 < 193 \cdot 1.56$). In case a larger statistics is needed to investigate the distribution of accidental triggers, when three or more sites are involved we can repeat the time slides on the same reference detector applying a fixed time offset to another site. We used this possibility for the V1H1L1 network: also the H1 detector has been shifted by a time of 1.56 s and 3.12 s. In this way we obtain three indipendent analyses composed of 192 lags each and we triplicate the livetime.

In the following considerations we indicate reference detector as the first in the name of each configuration. For instance, the reference detector for the network V1H1H2L1 is V1, for L1H1H2 is L1, and so on. We have not a complete freedom for the choice of reference detector: co-located detectors show a foreground noise made by correlated local disturbances, so they should not be shifted one with respect to the other. In fact, the existing foreground noie of H1 and H2 is a source of accidental triggers for the network configurations which include that pair of detectors, and shifting H1 with respect to H2 would eliminate this contribution. Another consequence of the H1H2 foreground noise is that time lag analysis cannot be used to measure the false alarm probability of H1H2 on-source candidates. Therefore, since we are lacking a robust estimate of its accidental background distribution, the H1H2 configuration has been analyzed only for setting a conservative upper limit on the GWs.

Background analyses have been used to tune the analysis thresholds. We decide to divide the background events list in two equal parts: the one belonging to the *odd* lags and the other belonging to the *even*. Only the first list has been used to tune the post production cuts (**tuning set**) at a chosen false alarm rate, while the second has been kept undisclosed until the tuning was frozen (**estimation set**). Then we apply post production cuts to the second set to estimate the actual false alarm rate of set by the chosen thresholds as well as the false alarm rate of any found candidate for each network configuration. In this way we avoid the risk of biding the false alarm estimate and can also verify the consistency between odd and even lags regarding the false alarm estimation.

For network configurations involving both V1 and L1, we decided to shift V1. This choice was made according to preliminary tests on JW1 period, which show a better efficiency at the same post-production thresholds when V1 is used a reference detector. This is related to the asymmetric location of the detectors involved in the network. Light travel time between detectors is almost equal considering the couples V1-L1 and V1-H1 (about 27 and 24 ms), while between Hanford and

Livingston it is significantly shorter (about 10 ms). The algorithm computes the likelihood over the sky in the time reference of the first detector, i.e. the data of the other detectors gets shifted by the light travel time corresponding to a grid of sky directions (App. A.1). When choosing V1 as a reference detector the pipeline reconstructs sky directions using time delays bounded within ± 27 and 24 ms, instead if we choose L1, one of the maximum time delays is significant smaller. The choice of V1 as reference detector then gives more possibilities to the algorithm to look for coherent signals, so that in the ideal case one would expect an higher accidental background and the same detection efficiency for real GW signals. It turns out, however, that this is not the case and that it is convenient to keep V1 as reference, since the background increase is negligible and instead the measured detection efficiency gets improved, as demonstrated in tests on the JW1 period. This is likely due to the approximations used within the algorithm to find candidate coherent signals.

Tuning procedures for all the network configurations (included H1H2, which is slightly different from the others) are explained in the Chapter 5.3.4.

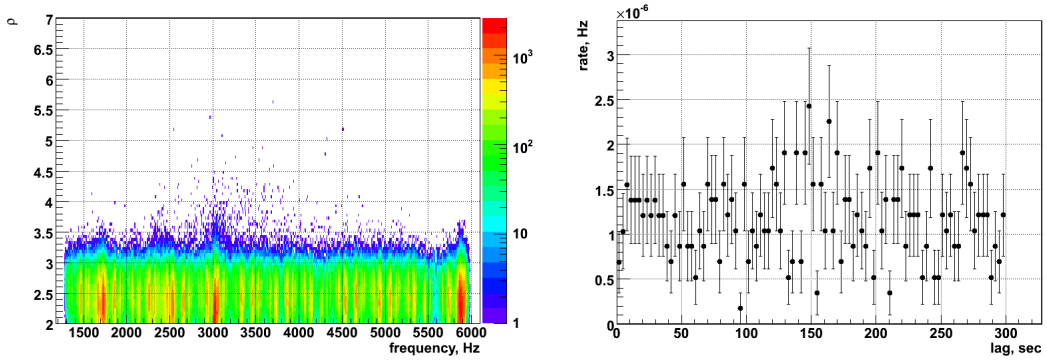

Figure 5.8: *Background behaviour of V1H1H2L1: ρ versus frequency (left) and rate versus lag (right). In this test, the post-production thresholds are set low enough to study also the background at low SNRs. The frequency dependence shows that accidental events are not white, but instead occur in frequency bands related to disturbances at different detectors. The overal count of accidental event per lag are however Poisson distributed and do not show a dependence on the time shift.*

5.3.3 Efficiency study

Detection efficiency depends on multiple factors: signal injection hrss, frequency, waveform, pipeline, in particular on the post production thresholds. We have just esplained the reasons on the choice to test CWB on SGQ9 set (Chapter 5.2.1).

This set is produced with a fixed hrss equal to $2.5 \cdot 10^{-21}$ $1/\sqrt{Hz}$. To test a wide range of signal amplitudes, CWB uses a grid of factors to scale the reference amplitude and repeat the software signal injections of the same MDC set. In this way, we construct an **efficiency curve** for each waveform, i.e. a plot showing the efficiency value as a function of injected hrss (or amplitude scaling factor). The range of these scaling fatcors should cover the entire detection efficiency curve, from near 0% to near 100% efficiency per each waveform.

We have chosen 18 amplitude factors, logaritmically spaced which allow this investigation for almost all tested waveforms (Fig. 5.9).

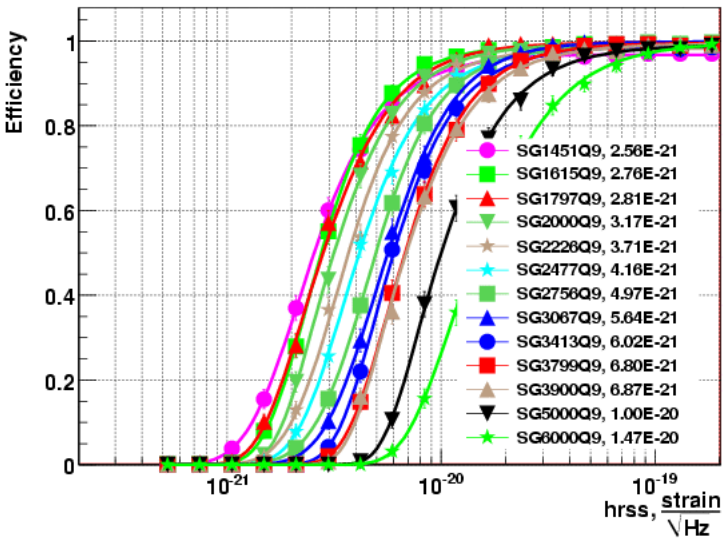

Figure 5.9: *Efficiency curves for SGQ9 set, V1H1H2L1 network, JW1 period. This is a test, post-production thresholds correspond to a higher false alarm rate with respect what is required for a GW search.*

From efficiency curves, we fit the values of injected hrss correponding to the 10%, 50% and 90% of efficiency and we consider these values as figure of merits among different waveforms or searches.

Efficiency curves also guide the astrophysical implementation of the search results.

5.3.4 Thresholds tuning

The main goal of the tuning procedure is to select events making a balance between the probability of false alarm and the detection efficiency.

The adopted tuning procedure is the following:

- target a false alarm rate suitable for foreground investigations and for setting upper limits, which is approximately $5 \cdot 10^{-9}$ 1/s ($\sim$1 over 6 years);
- select post production thresholds so to maximize the detection efficiency for SGQ9 keeping approximately constant the target FAR;
- estimate the false alarm rate of events passing the threshold of previous step by a procedure defined a priori, which ranks the significance of selected events as a function of their ρ. The number of time lags will be increased as needed to ensure the statistical uncertatinty on FA.

We remind here the post production thresholds (Chapter 3.3):

- Penalty: $p > p_{th}$
- Effective Correlated SNR: $\rho > \rho_{th}$
- Network Correlation Coefficient: $cc > cc_{th}$

in addition we have added a frequency cut on networks involving H1 to consider only events with central frequency outside the band [5482.5 - 2.5, 5482.5 + 2.5] Hz for calibration problems on H1 (Chapter 5.1.2). Events which satisfy the condition $|f - 5482.5\text{Hz}| > 2.5$ Hz pass this cut (f is the estimated central frequency).

We remind that **penalty cut** (Sec 3.3.1) is is defined as the square root of the fraction between the total energy in the detector and the reconstructed energy response and it is used to avoid unphysical solutions when energy response is greater than total energy in the detector data stream. This is a general condition and it is not necessary to tune it for the different networks, so we decide to adopt the threshold used in the LF frequency, which has been previously tested in the LF band. We do not expect relevant differences regarding this cut. The p_{th} is set equal to 0.6 for all networks.

The other two thresholds, instead, have to be tuned depending on the performances of the detector configurations. Results performances gets better when the number of detectors involved is higher, because the network gives more information to the pipeline. We cannot use the same post production cuts of the LF search, because noise behaviour between detectors on HF is different from LF.

We adopt a common procedure to choose the values of ρ_{th} and cc_{th} for each configuration, and we apply this procedure on **tuning set** (odd lag) and SGQ9 simulations.

First we consider the ρ versus cc distribution of background and SGQ9 events (Fig. 5.10). From this plot we select some convenient cc thresholds (usually

4 thresholds with a coarse step of 0.5) which can separate the populations of reconstructed injections and of accidental triggers. Referring to Fig. 5.10, cc_{th} too near to unity would eliminate too many SGQ9 events, otherwise for very low values of cc_{th} too many background events would pass.

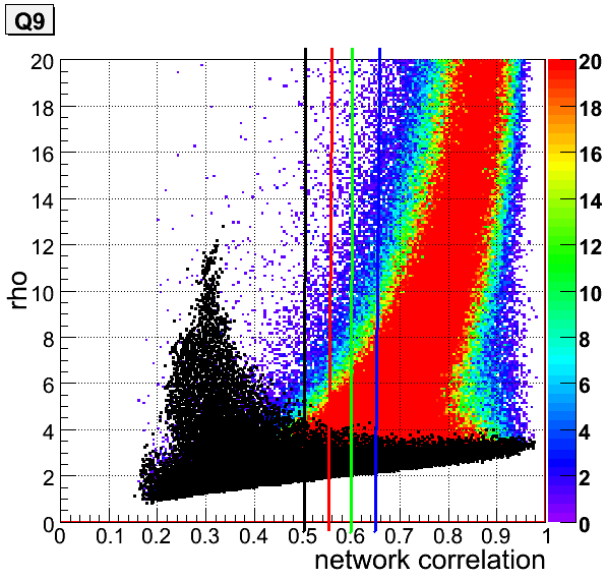

Figure 5.10: *Distribution of ρ and cc values for background (black) and SGQ9 (coloured) events for V1H1H2L1 network. Coloured lines show candidate thresholds on cc: black→0.5, red→0.55, green→0.6, blue→0.65.*

For each value of *cc* thresholds chosen we produce the accidental event rate versus ρ (Fig. 5.11) and we find the corresponding ρ_{th} ensuring the target false alarm rate.

We compare the resulting hrss 90% detection efficiencies of these different [ρ, *cc*] settings for the SGQ9 set (Fig. 5.12) and select the one that performs best in the band 2-6 kHz.

Final thresholds are selected automatically using an implemented procedure. This ranking procedure scores more the lowest hrss 90%. It considers the couples [ρ_i, cc_i] for each network i. Then, for each SGQ9 frequency, we have hrss 90% values corresponding to each couple: hrss90%[1] and hrss90%[2]. The procedure assigns a score to the couple [ρ_i, cc_i] that at each frequency beats by at least 10% the hrss90% of the others. After calculating the scores for each [ρ_i, cc_i] couple, the procedure selects the couple with higher score. If two or more couples have the same score, the procedure choose the one with lowest *cc* value.

For the **V1H1L1** network a difference between tuning and estimation sets of

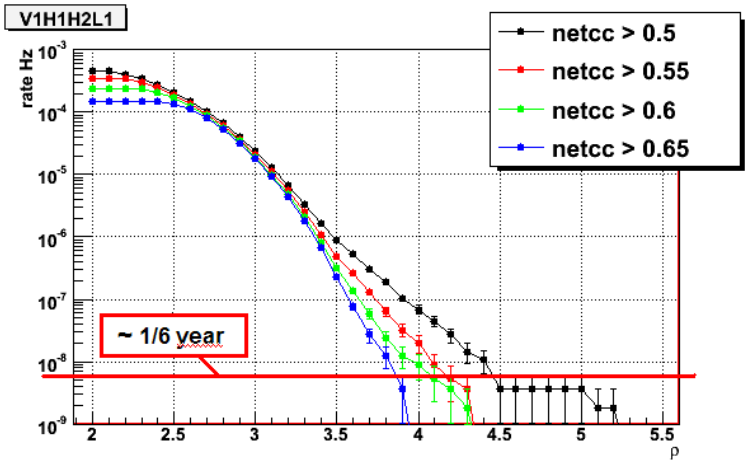

Figure 5.11: *Accidental event rate versus ρ for V1H1H2L1 network. The red line shows the target false alarm rate of ∼1/6 years.*

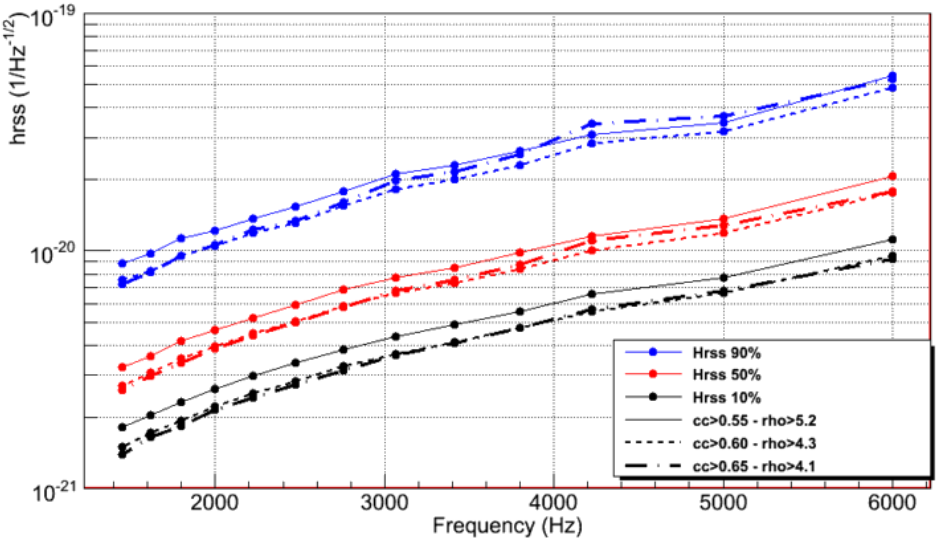

Figure 5.12: *Hrss 10-50-90% efficiency for SGQ9, V1H1H2L1 network, per each pair of (ρ, cc) thresholds which match the target false alarm rate.*

the background forced us to set more conservative thresholds with the respect to the automatic procedure. This is described in section 5.5.2.

Final thresholds for all configurations are on Tab. 5.8

Network	$\mathbf{p_{th}}$	$\mathbf{cc_{th}}$	$\rho_{\mathbf{th}}$
V1H1H2L1	0.6	0.6	4.3
L1H1H2	0.6	0.6	4.1
V1H1H2	0.6	0.7	4.6
V1H1L1	0.6	0.6	6.0
H1H2	0.6	0.5	5.0

Table 5.8: *Final values of p_{th}, ρ_{th} and cc_{th} for the different network configurations.*

H1H2 Tuning

The tuning of H1H2 configuration cannot be performed in the same way as the other configurations. This is due to the fact that we cannot characterize H1H2 background with time lags between the two detectors, as we have previously decided. We decide to make statistical analysis on H1H2 configuration with the goal to compute upper limits on GW, so the tuning procedure is slightly different: given the unknown foreground systematics, we select more conservative thresholds with respect to the accidental event distribution obtained from time lags.

We decide to characterize background with different H1H2 analyses:

- **zero lag playground**: zero lag analysis on time periods corresponding to V1H1H2L1, L1H1H2 and V1H1H2 network observative time (approximative 208 days). Since a null results has been achieved for these configurations, this is almost an off-source sample for GWs. We are not assured of a null result for H1H2, but it is very likely.

- **off-source playground**: 192 lag analysis on time periods corresponding to V1H1H2L1, L1H1H2 and V1H1H2 network observative time. This is an off-source sample, but any foreground noise due to correlated H1H2 responses is canceled as well

- **off-source background**: 192 time lags on the H1H2 exclusive live time.

We select final thresholds in the following way. Looking at the scatter plot of accidental events and MDC SGQ9 injections (Fig. 5.13), we fix a priori the *cc* threshold equal to 0.5. In fact, the ρ threhold is set extrapolating above the loudest event among the three analysis considered (off-source background, Fig. 5.14). We consider both the odd and even lags (no separation between tuning and estimation set) for each of the three analysis to have more statistics, in fact the number of background events is very small, respect to the other configurations of detectors (Figs. 5.12 and 5.13). This is due to the fact that they are colocated, so that the expected coherent response in the two detectors is so welle constrained

that makes much less likely for accidental triggers and for foreground triggers to pass the coherent analysis and conseguently the reconstructed background events are reduced.

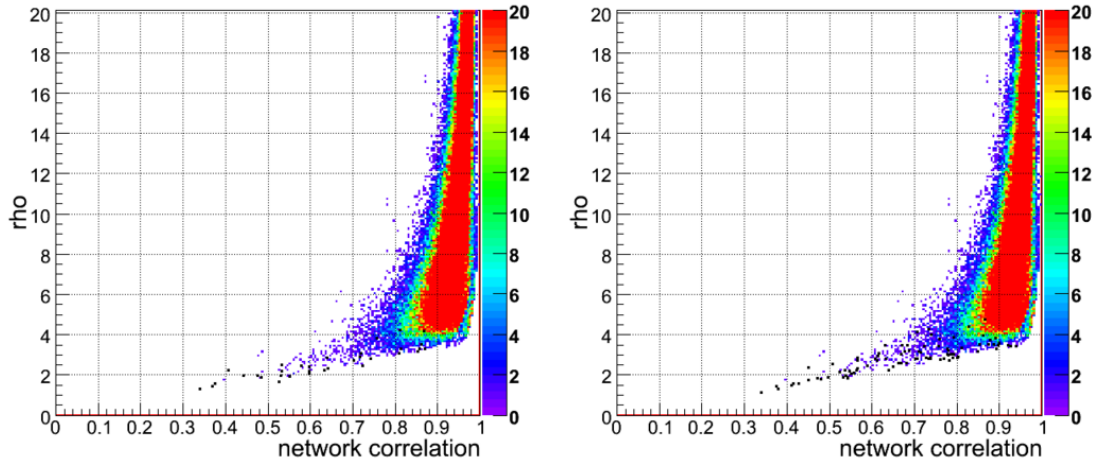

Figure 5.13: *Distribution of ρ and cc values for background (black) and SGQ9 (coloured) events for H1H2 network. Off-source background on left and off-source playground on right.*

Comparing SGQ9 efficiency results for different ρ thresholds (Fig. 5.15) we realize that the choice of ρ_{th} can be conservative, without paying too much on the efficiency side. The loss in efficiency is very small increasing ρ threshold from 4 to 5, so we can choose the most conservative value, i.e. above the loudest event, for our goal. So the chosen ρ_{th} is equal to 5.0.

5.4 Effects of Calibration Systematics on results

To test the robustness of our detection efficiency results with respect to calibration systematic uncertainties, we performed MonteCarlo software injections using ad hoc mis-calibrated signals. The difference in hrss 50% between calibrated and mis-calibrated data can be used as an error of the efficiency estimate due to calibration uncertainties.

The pipeline has been sligthly modified to perform mis-calibration tests. For amplitude/phase mis-calibrations, before injecting MDC waveforms in the data stream, we modify on the fly their amplitudes/phases according to the available estimate of calibration systematics. Usually we indicate applied mis-calibrations

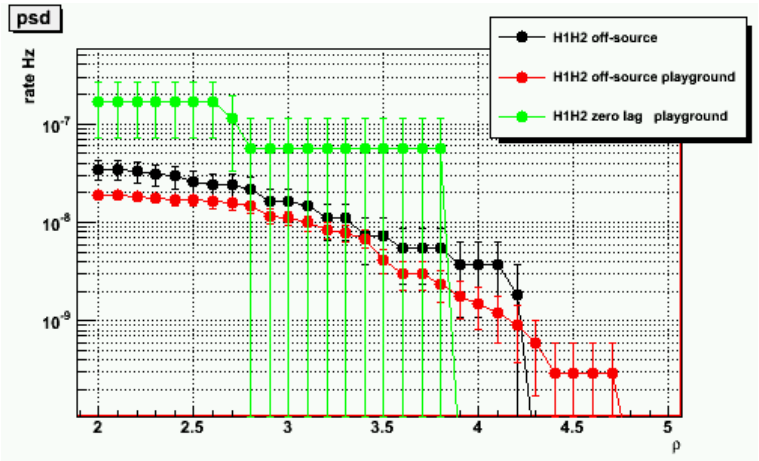

Figure 5.14: *Accidental event rate versus ρ for H1H2 network and off-source background (black), off-source playground (red), zero lag playground (green). Zero lag playground has a very porr statistics but anyway does not contradict the other estimated. The off-source playground and background are in agreement in the tail and the former corresponds to a much longer livetime and therefore can investigate lower False Alarm Rates.*

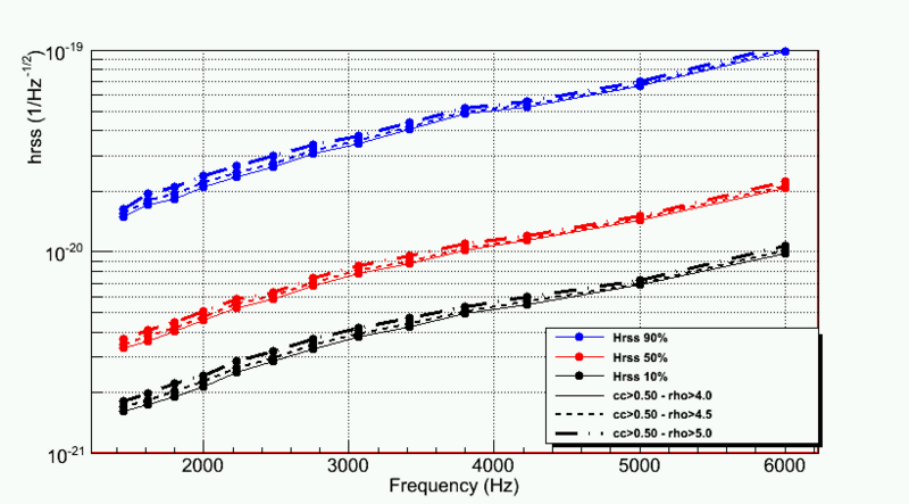

Figure 5.15: *Hrss 10-50-90% efficiency for SGQ9, H1H2 network, obtained using different ρ thresholds and the same cc.*

with the relative fraction respect to the original amplitude or with the equivalent number of samples for the phase.

As a matter of fact, phase shifts are technically implemented by delaying or

anticipating the waveform by an integer number of samples. Minimum time shift depends on the sample rate. Starting from a sample rate equal to 16384 Hz[4], the minimum time shift applcable would be about 60 μs, which corresponds to about 40 degrees at 2 kHz and 120 degrees at 6 kHz. We considered this too coarse, so we up-converted MDC stream to a sample rate of 4*16384, in this way the minimum time shift is about 15 μs.

These tests are performed on JW1 period and mis-calibrations values depend on calibrations version Estimated errors are in the Tab. 5.5

Amplitude mis-calibrations tested are of $\pm$ 10% for the several detectors in the different networks. We considered a maximum phase difference of 32° at 6kHz between LSC detectors, which correspond to a time shift of 15 μs. To test the effect of much greater systematics in between LSC detectors and Virgo, we considered a delay time of 180 μs between V1 and L1. Amplitude and Phase shifts applied are listed in the Tab. 5.9

Results of these tests show the robustness of the algorithm respect to mis-calibrations. Significative loss on efficiency are present when co-located detectors are mis-calibrated to each other, especially for the phase shift (Fig. 5.17). This is due to the fact that introducing a time shift for co-located detectors their response become incherent, being out of phase.

These tests are useful to test the robustness of the algorithm, but they are not sufficient for the estimation of efficiency results. For the calculation of upper limit, in fact, we have to take in account the estimation error due to mis-calibration effects. For this purpose, we consider the calibration errors at each frequency due to LSC-V4 and Virgo-V2 uncertainties and systematic differences between LSC-V3 and V4. The former can be modeled as one sigma uncertatinties (estimated by rms of calibration results) while the latter is a mean of systematic deviation (Tab. 5.5).

We developed a new procedure to estimate the effects of the known calibration systematic uncertainties on the detection efficiency of each network configuration. This procedure aims at folding the effects of calibration systematics in the measurement of the detection efficiency. This is accomplished by sampling the amplitude of each software signal injection of the SGQ9 MDC set from a Gaussian distribution, whose mean and sigma are the best estimate of the amplitude calibration deviation and its rms fluctuations at each detector (Tab 5.5). In this way, per each SGQ9 central frequency, we get a detection efficiency measurement of each network configuration which includes the effects from known calibration systematics.

To have an actual estimate of the uncertainty of the resulting detection efficiency, this procedure was repeated more times, each time sampling only once the amplitude deviation for all injections from the assumed gaussian distribution

[4]Mis-calibrations are implemented before Down Conversion.

Network	**Amplitude/Phase**		**Detectors**			
			L1	**H1**	**H2**	**V1**
V1H1H2L1	Amplitude (%)		+10	-10	0	0
	Phase	Sample	12	12	11	0
		Time (μs)	183	183	168	0
		Degrees at 2 kHz (°)	132	132	123	0
		Degrees at 6 kHz (°)	395	395	371	0
L1H1H2	Amplitude (%)		+10	-10	0	
	Phase	Sample	1	1	0	
		Time (μs)	15	15	0	
		Degrees at 2 kHz (°)	11	11	0	
		Degrees at 6 kHz (°)	32	32	0	
V1H1H2	Amplitude (%)			-10	0	0
	Phase	Sample		12	11	0
		Time (μs)		183	172	0
		Degrees at 2 kHz (°)		132	123	0
		Degrees at 6 kHz (°)		395	371	0
V1H1L1	Amplitude (%)		+10	-10		0
	Phase	Sample	12	12		0
		Time (μs)	183	183		0
		Degrees at 2 kHz (°)	132	132		0
		Degrees at 6 kHz (°)	395	395		0
H1H2	Amplitude (%)			-10	0	
	Phase	Sample		2	1	
		Time (μs)		30	15	
		Degrees at 2 kHz (°)		22	11	
		Degrees at 6 kHz (°)		65	32	

Table 5.9: *Amplitude and Phase mis-calibrations tested for the different networks.*

of systematic uncertainties. The rms fluctuations of these different detection efficiency are then estimates of one sigma uncertainties on the detection efficiency due to amplitude calibration uncertainties.

The assumption of Gaussian distribution is coarse and it is good only for 1σ estimates. However, this procedure is particularly well suited to fold calibration systematics in the results of coherent searches. In fact, in coherent searches it is difficult to select a priori which micalibration conditions may give worst results: in addition to the naive worsening due to a systematic underestimation of the h(t) data streams a coherent algorithm sensitive also to mis-balance in calibrations among detectors.

The actual MDC simulations have been performed on the JW1 data set to save computational time, since detection efficiency of JW1 data is representative of all the S5Y2/VSR1 period.

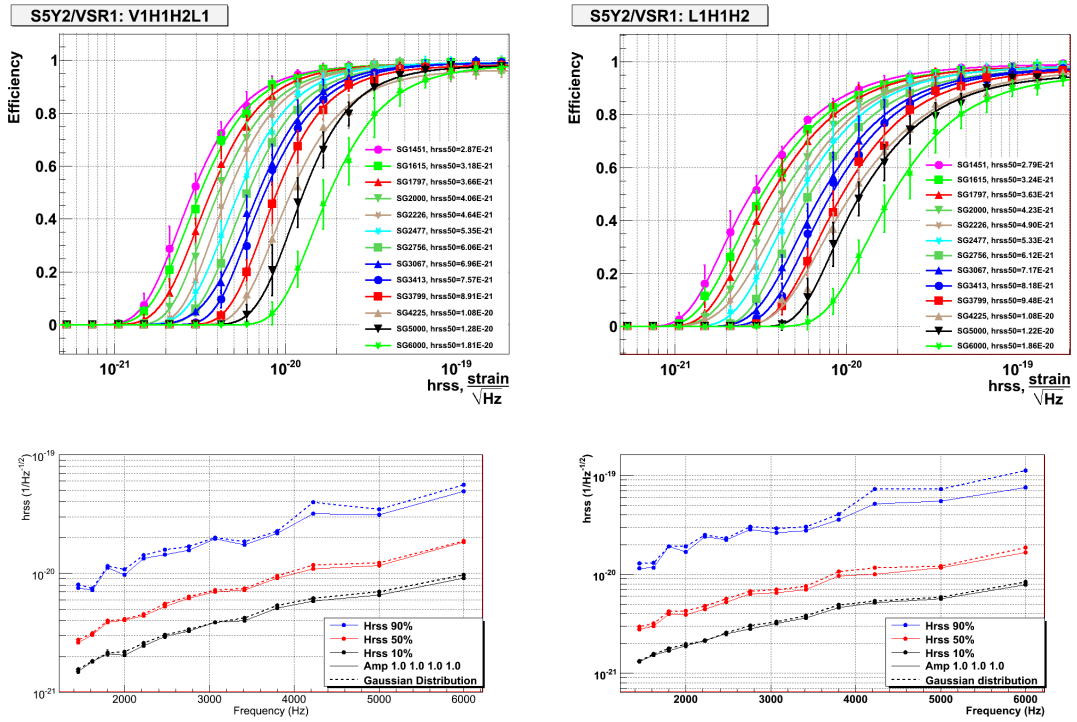

Figure 5.16: *Efficiency curves (top) and Hrss10-50-90% (bottom) resulting from Gaussian mis-calibrated tests. The error bars on efficiency are calculated from analyses made selecting randomly 18 values of mis-calibrations sampled from the Gaussian distribution. V1H1H2L1 on the left and L1H1H2 on the right.*

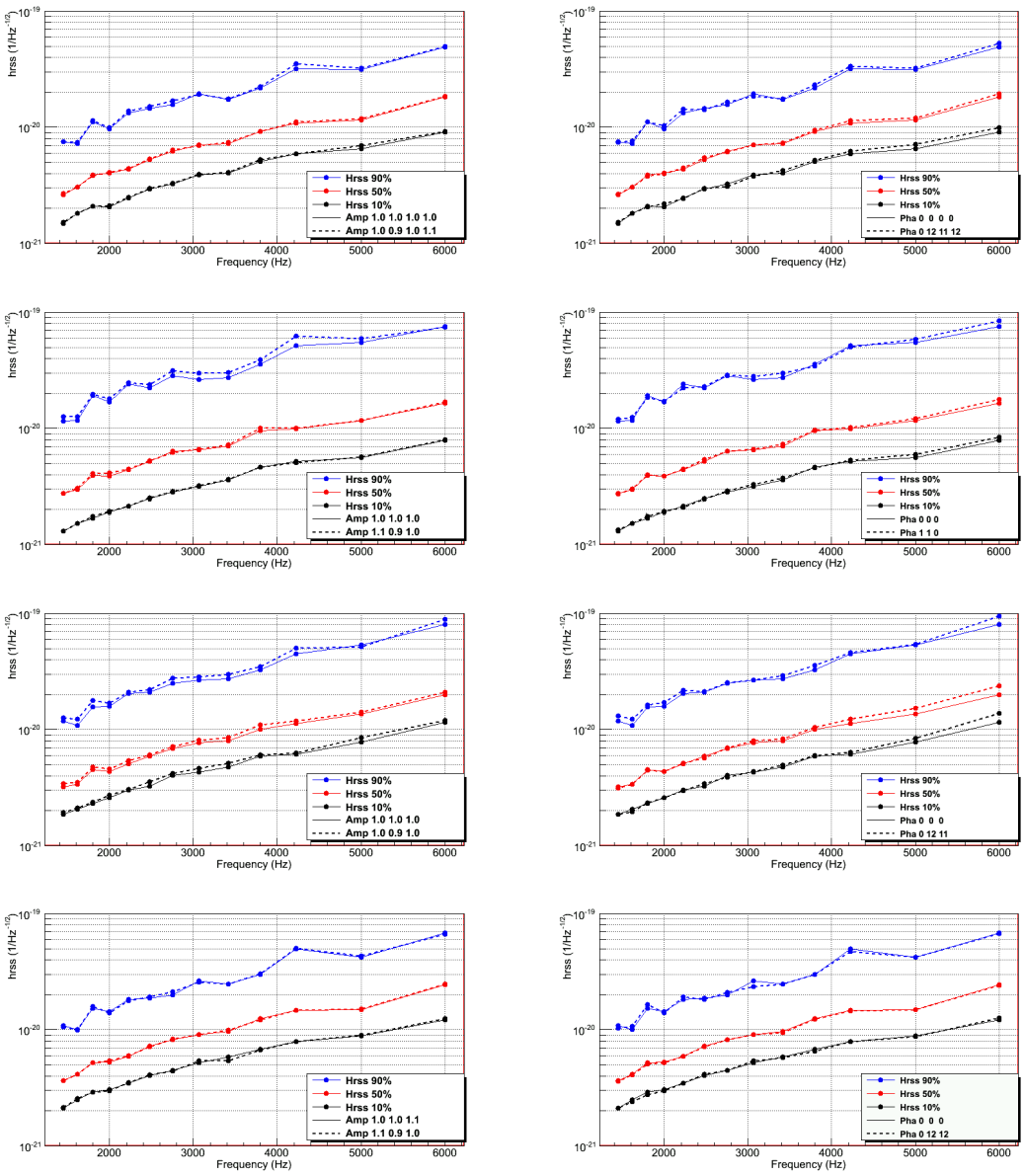

Figure 5.17: *Hrss 10-50-90% for selected amplitude (left) and phase (right) miscalibration effects. From top to bottom: V1H1H2L1, L1H1H2, V1H1H2, V1H1L1, H1H2.*

5.5 Sanity and Consistency checks

Checks on results are useful to verify the correctness of work. A comparison between odd and even lags, for instance, gives an idea on the quality to the

choice of post-production thresholds (**consistency**). Other checks (loud injections, events reconstructed outside Category II segments) can test for bugs in the pipeline (**sanity**).

5.5.1 Events reconstructed outside Category II segments

Central time of reconstructed events should occur within the useful science data (Category II segments), otherwise this could be an alarm for bugs in the pipeline. This check has been performed both for background and SGQ9 simulations. We found a small fraction of events oustide the Category II segments (Tab. 5.10). Analysing the reconstructed times of these events, we found that their differences between the start (or stop) of the nearest segment is less than 20 ms. This is consistent with the CWB method to select time periods. The algoritm selects time (and also frequency) making boundary cuts on time-frequency planes, after the application of the Wavelet Transform. This selection is not exactly applied to the desired time, but it is approximated to the time resolution of the wavelet decomposition level at which we apply the boundary cuts. Starting from a sample rate of 12800 Hz and using from 4 to 9 decomposition levels, the maximum time resolution is about 40 ms. So we can expect a small fraction of events with a central time outside of the segments less than half this value this value. This because pixel time is centered on the pixel.

Network	**Background**		**Simulations**	
	Total	Outside	Total	Outside
V1H1H2L1	5857829	218	803845	182
L1H1H2	1134043	3	1223954	12
V1H1H2	398698	4	165665	58
V1H1L1	2201361	117	61358	25
H1H2	504	0	288947	0

Table 5.10: *Number of events with a central time outside Category II for Background and SGQ9 simulations.*

5.5.2 Consistency between odd and even lags

The tuning and estimation sets for each network should be two indipendent samples of the accidental background distribution. We use the tuning set to choose post-production thresholds and the estimation set to calculate false alarms avoiding any possible statistical biases introduced by the tuning process. Before applying

post-production thresholds to the zero lag, we verify if the statistical properties of the two samples match for all the networks (excluding H1H2 network). Fig. 5.18 shows the comparison between odd (black) and even (red) false alarm rate as a function of ρ_{th}. Rate behaviour is compatible between odd and even lags for all networks except for the V1H1L1 one.

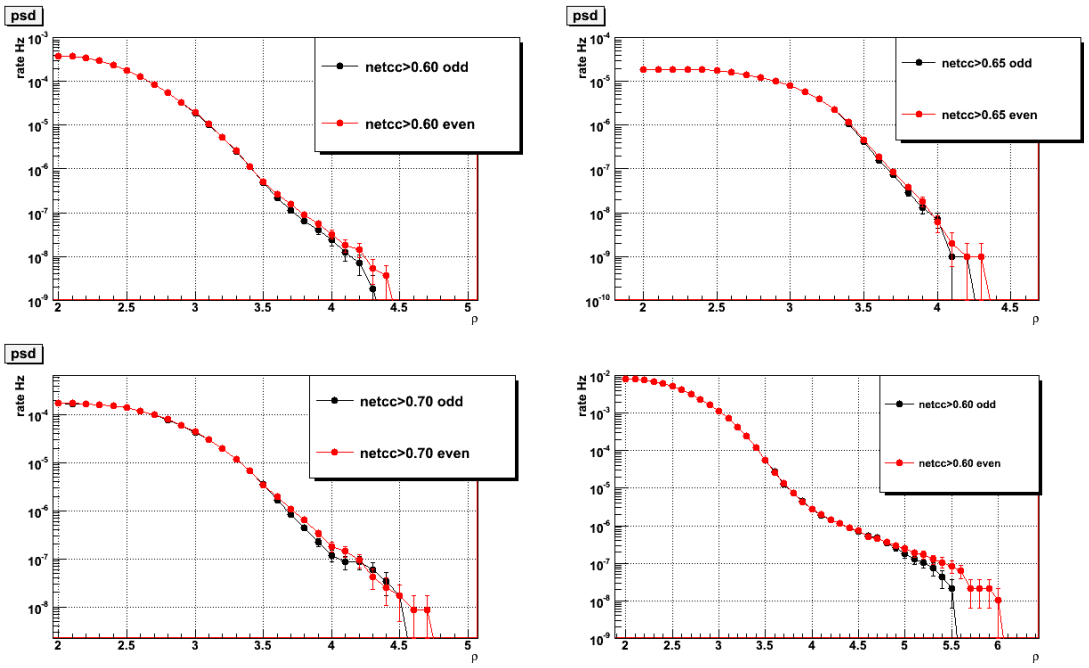

Figure 5.18: *Consistency between odd (black) and even (red) lags background rate as a function of ρ_{th} for different newtork. From left to right and tob to bottom: V1H1H2L1, L1H1H2, V1H1H2, V1H1L1.*

We investigate the cause of this behaviour and we find a bump of noisy events with high ρ and cc (Fig. 5.19, left). These events are well localized in time (Fig. 5.19, right), so we suspect a non standard behaviour of some instruments not spotted by the Data Quality cuts. This inconsistency between odd and even lags for this configuration is an evidence that the statistical uncertainties estimated for the odd and even sets are underestimated. So we decided to set more restrictive thresholds using also the information from the even lag set, which dominates the empirical distribution of the loudest accidental events.

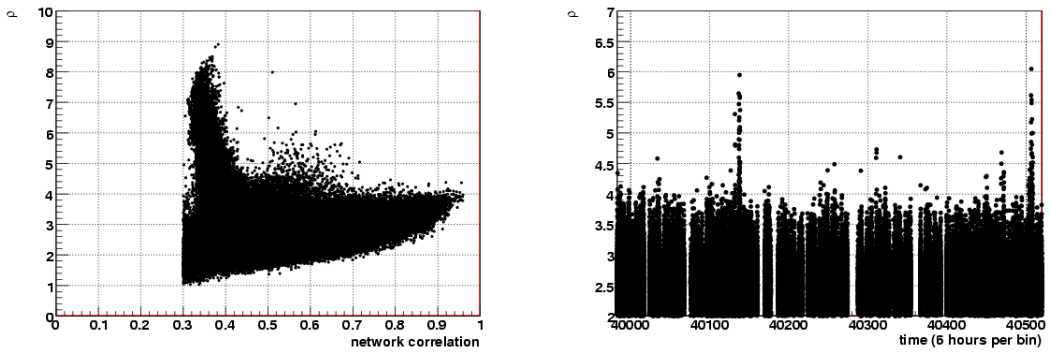

Figure 5.19: *Scatter plot of ρ and cc (left) and ρ versus time (right) for the V1H1L1 estimation set. The plot on the left reporst events with* $cc > 0.5$.

5.5.3 Missing Loud Injections

We call "Loud Injections" the ones with the highest h_{rss} amplitude at earth, corresponding to the largest amplitude factor of the MDC injections. In fact, given their high h_{rss} level, we espect that these loud injections should be correctly identified by the analysis, with a few exceptions in case they occur from polarization and sky locations which show a very unfavorable antenna pattern.

Simulation tests show a low number of missing lound injections. There are two ways which a Loud Injections can be missed: either the algorithm fails to reconstruct it, or it misses it because it not pass the post-production thresholds (Tab. 5.11).

Network	**Injected**	**Found**	**Not Pass**	**Not Rec**
V1H1H2L1	58613	58024	563	26
		99%	0.96%	0.04%
L1H1H2	104523	102356	1867	300
		97.9%	1.8%	0.3%

Table 5.11: *Loud Injections for V1H1H2L1 and L1H1H2: Injected, Found, Reconstructed but not passing Thresholds (Not Pass), Not Reconstructed (Not Rec).*

We can verify that the events that are not reconstructed by the algorithm have unfavourable antenna patterns. Whereas, events not passing post-production thresholds shows an high disbalance among reconstructed signal energies in different detectors. Fig. 5.20 report on the left the value of $|F_\times|/|F_+|$ in the Dominant Polarization Frame: events that are not reconstructed have source sky positions

where this value is almost null. On the right there is a scatter plot of V1 and H1 SNRs. Black points are reconstructed events passing threholds, instead red points are reconstructed events not passing thresholds. This last group shows an high disbalance on detectors SNR values.

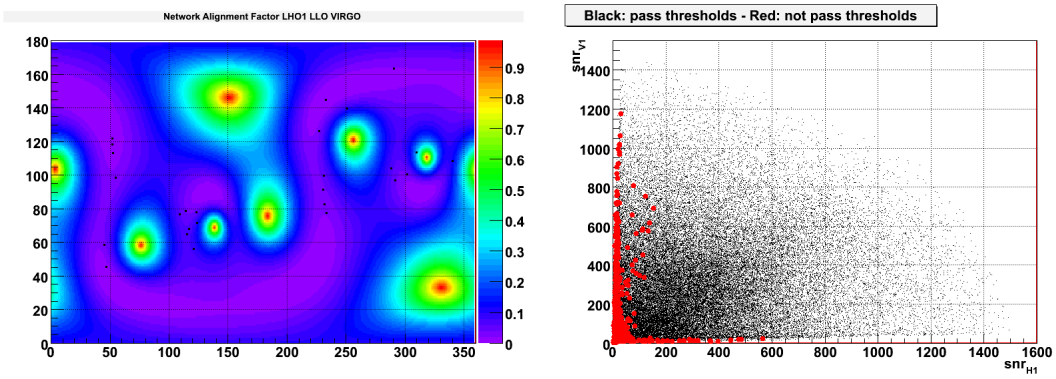

Figure 5.20: *Left: Fraction of plus and cross antenna pattern defined in the DPF. Black dots are not reconstructed events. Right: V1 SNR versus H1 SNR for reconstructed events passing (black) and not passing (red) thresholds.*

5.5.4 Comparison with LF search

Low and High Frequency analyses have an overlapping frequency band between 1280 and 2048 Hz. This overlapping band can be used to verify the consistency between the two analyses. We remind that post-production thresholds are not equal between LF and HF searches, because sensitivity curves and glitches depend on frequency (LF search has performed two separate tunings for frequency band [0Hz, 200Hz] and [200Hz, 2048Hz]). Discrepancies in the final detection efficiencies for waveforms in the common bandwidth should be small, given the similar choice of post production thresholds. For SGQ9 we have two common waveforms in this frequency band: SG1615Q9 and SG2000Q9. We report in the Tab. 5.12 the comparison of Hrss50% and Hrss90% for each network, values are compatible. However, we expect more discrepancies on the SG2000Q9 waveform, due to the frequency band choice on the LF: [64, 2048] Hz. In fact, because of this frequency band selection, about a 40% of the signal is lost (without taking in account the noise). This could causes a loss on efficiency, especially for less energetic signals (lower values of injected h_{rss}). So, we expect worse values of Hrss50% and Hrss90% for the LF results, as it is on Tab. 5.12.

Network	**Waveform**	**LF** (10^{-21} $1/\sqrt{Hz}$)		**HF** (10^{-21} $1/\sqrt{Hz}$)	
		Hrss50%	Hrss90%	Hrss50%	Hrsss90%
V1H1H2L1	SG1615Q9	3.08	8.51	3.07	8.26
	SG2000Q9	4.43	12.4	3.94	10.4
L1H1H2	SG1615Q9	3.32	14.9	3.88	16.8
	SG2000Q9	4.76	20.5	3.88	16.8
V1H1H2	SG1615Q9	3.18	14.6	3.55	12.6
	SG2000Q9	5.40	22.5	4.42	17.5
V1H1L1	SG1615Q9	4.62	1.47	4.76	1.74
	SG2000Q9	5.74	1.63	5.93	1.57
H1H2	SG1615Q9	3.60	17.8	4.11	19.6
	SG2000Q9	5.02	23.7	5.08	23.7

Table 5.12: *Comparison between Hrss50% and Hrss90% for LF and HF searches. The SG2000Q9 shows in general larger differences0dip, due to the fact the for the LF it is applied a frequency band selection on [64, 2048] which cuts about 40% of the signal.*

5.6 Results

Once chosen the post-production thresholds for each network and performed the above checks on the analysis, we are ready to analyze the on-source data. For each network we apply post-production thresholds to zero lag triggers and we verify it there are candidate surviving the thresholds.

In case candidate events are found above the final thresholds, they would be considered for further investigations. In particular their false alarm rate would be estimated against the empirical distribution of accidental events by ranking according to their ρ, and eventually excluding to a detection check list (DQ Category III, see Sec. 5.1.1) whose aim is to perform deeper sanity checks and provide a posteriori information on the characteristics of the detectors and of the reconstructed signal which could help in assessing an a posteriori confidence or plausibility for a GW or mondane nature of the candidate. In fact, the statistical study esplained above is not able to discriminate the nature of selected candidates, in particular a GW signal or a foreground event, the detection check list should help to shed light on the physical source of the candidate signal.

Moreover, the candidate is checked with the "Blind Injection", which is a list of hardware injections not considered by DQ. These injections times are not unveiled until the boxes are opened and they are used as a check on the algorithms performances. Blind injections have the same characteristics of an expected GW signal, so they should be detected and considered candidate events if strong enough, unless they occur at times later excluded by DQ vetoes. Unfortunately, when they are

injected, DQ segments are not ready, so it is possible that the blind injections are excluded by the DQ vetos, and so they are not reconstructed by each algorithm.

If no events pass post-production thresholds, or if no candidates are accepted as GW detection candidates, the results of the counting experiment will be used to set *upper limits* on the rate on GW events. This is accomplished by setting an upper limit to the counting experiment according to the unified approach by Feldman and Cousins [113]. Then this upper limit is interpreted in terms of rate versus signal strength at earth per each waveform by taking into account the measured detection efficiency and the observation time. Here the rate of GWs is intended as the random rate of a Poisson point process.

For the five networks analysed, we have **null results** on zero lag, i.e. no candidate events. Fig. 5.21 shows the comparison between rate as a function of ρ_{th} for background (black) and zero lag (red). The distributions of events rate versus rho at zero lag are in agreement with the measured distributions of accidental events within the expected statistical uncertainties. The events here have been selected above the cc_{th} and the comparison have been performed, using all reconstructed triggers, below the ρ_{th} threshold. The zero lags appear therefore just as samples of the expected accidental distributions. Moreover, a null results in all the trials is in agreement with the overall false alarm probability (FAP) of the search, which is about $FAP \approx 7.089 \cdot 10^{-2}$.

In fig. 5.22 there are the efficiency curves for SGQ9 waveforms. These figures are calculated at the final post-production thresholds.

From these results we can calculate **upper limits** for the injected waveforms. Upper limits is a confidence interval: in particular the upper limit will ensure a chosen conservative coverage, i.e. results would surpass by chance the upper limit with probability less or equal to 1-coverage when only accidental events are present.

Upper limit can be easily calculated from a single pipeline considering the rate of a Poisson-distributed process. We consider an experiment that measures the number of events of a specific random process that occurs in a time T. We introduce the following quantities: a mean rate μ of foreground events (that occur independently from each others and is unknown); the efficiency ϵ of successfully detecting any given event; a mean number b of background events in the time T. We assume that the total number of events (foreground plus background) counted in a given time is Poisson distributed. So, the probability of detecting N events over the full period T is given by:

$$P(N|\epsilon\lambda + b) = \left(\frac{(\epsilon\lambda + b)^N}{N!} \right) e^{-(\epsilon\lambda + b)} \tag{5.4}$$

where we have defined $\lambda = \mu T$ as the expected mean number of foreground events.

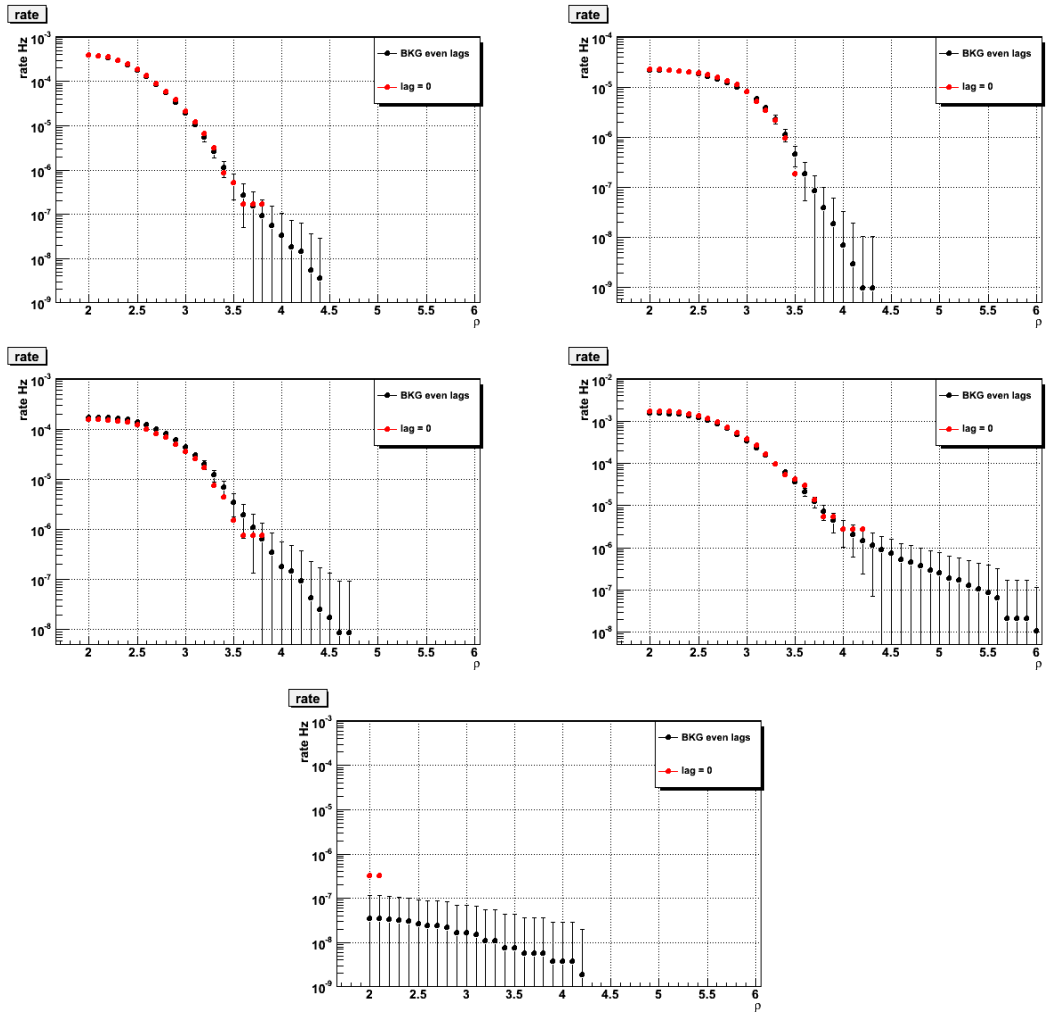

Figure 5.21: *Rate versus ρ_{th} at the final cc_{th} for the network. From left to righ and top to bottom: V1H1H2L1, L1H1H2, V1H1H2, V1H1L1, H1H2.*

The Feldman-Cousins approach [113] is to define a confidence interval $[\mu_1, \mu_2]$ which is a member of a set, such that set has the poperty that:

$$P(\mu) \in [\mu_1, \mu_2] = \alpha \tag{5.5}$$

Here, μ_1 and μ_2 are functions of the measured variable and the Eq. 5.5 refers to the varying confidence intervals from an ensemble of experiments with fixed μ. For a set of confidence intervals, Eq. 5.5 is true for every allowed μ. If the Eq. 5.5 is satysfied, then we say that the intervals covers μ at the stated confidence, or equivalently, that the set of intervals has the correct coverage.

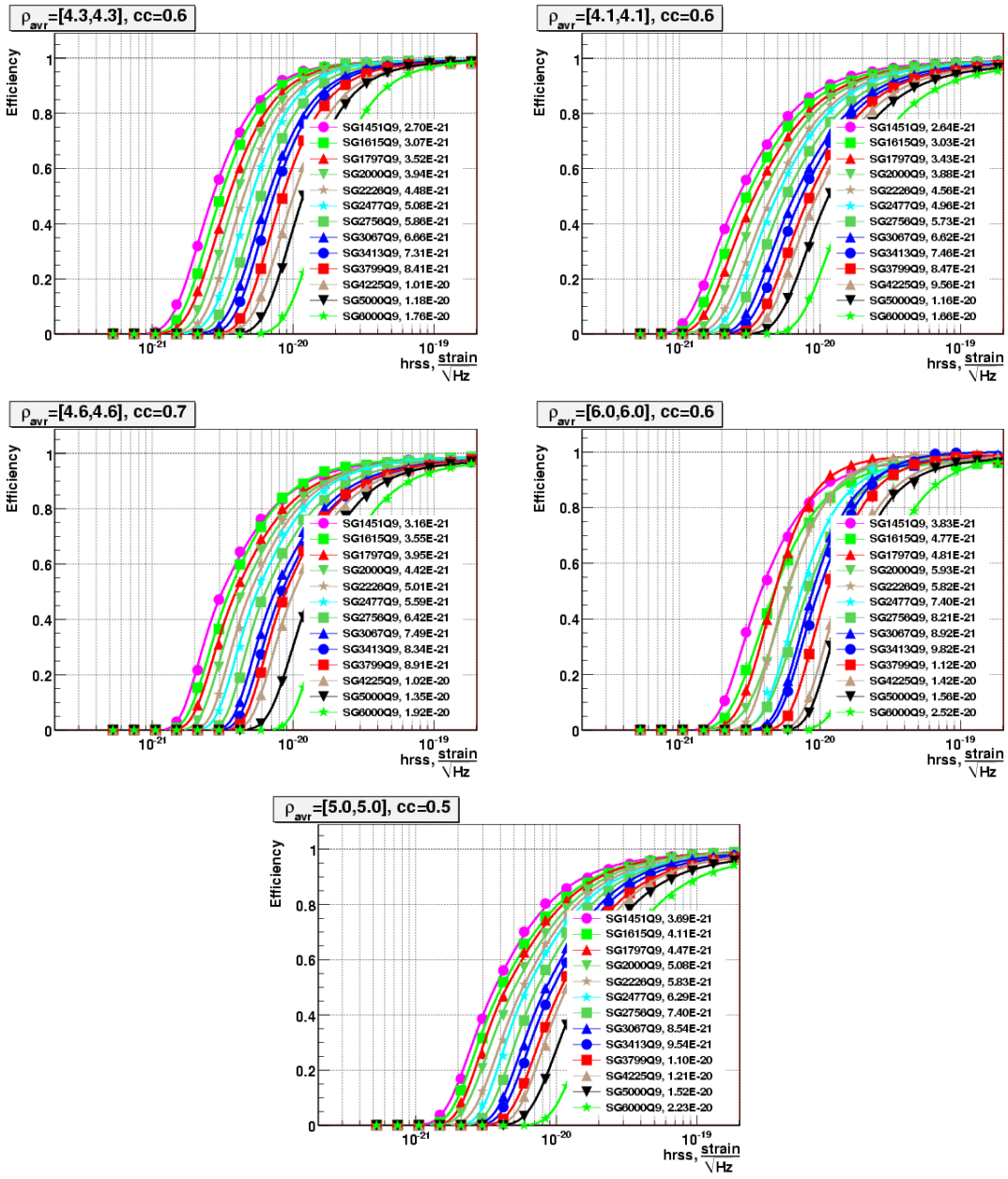

Figure 5.22: *Efficiency curves at final ρ_{th} and cc for the network. From left to righ and top to bottom: V1H1H2L1, L1H1H2, V1H1H2, V1H1L1, H1H2. In the table of each figure are reported the Hrss50% values for each waveforms.*

So, selected a confidence level α and given n measured events, the frequentist upper limit λ_α at a confidence level α is the value at which there is a probability

$1-\alpha$ of measuring more than n events.

$$1-\alpha = \sum_{N=n+1}^{\infty} P(N|\epsilon\lambda_\alpha + b) = 1 - \sum_{N=0}^{n} P(N|\epsilon\lambda_\alpha + b) \tag{5.6}$$

For example, if we want the 90% confidence level when we have not observed any events: $\alpha = 0.9$, $n = 0$ and $b = 0$. Inserting this value in Eq. 5.6 we obtain:

$$1-\alpha = 0.1 = P(0|\epsilon\lambda_{90\%}) = e^{-\epsilon\lambda_{90\%}} \tag{5.7}$$

and we calculate the value of $\mu_{90\%} = \lambda_{90\%}/T$:

$$\mu_{90\%} = \frac{2.30}{T\epsilon} \tag{5.8}$$

From this equation we can obtain upper limit for each single network analysis. An example is reported in Fig. 5.23 for the V1H1H2L1 network and SGQ9 waveforms.

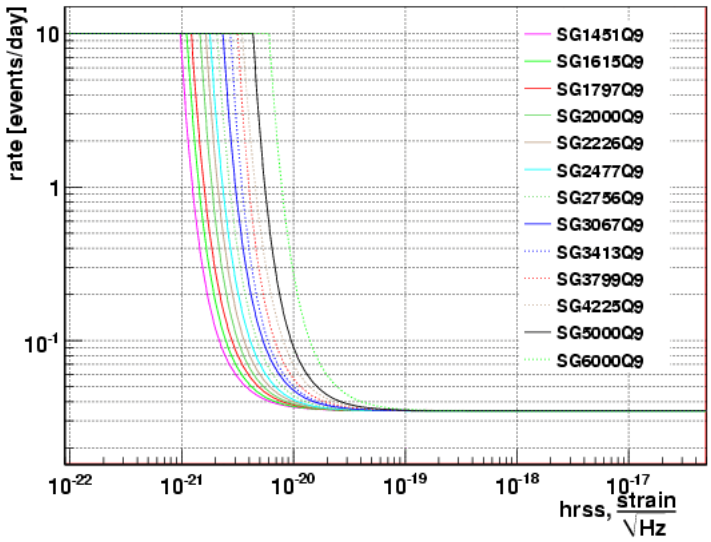

Figure 5.23: *Upper limit curves for SGQ9 waveforms calculated from V1H1H2L1 analysis using a 90% coverage.*

We use the approach described in [114] to combine the results of the difference searches, which combine sets of triggers according to which network detected any given trigger. This is a simplified approach with respect to the Feldman and Cousins. Considering two searches (A, B), we have for each quantities previously

introduced three subsets: A, B, AB. For instance: n_A is the number of events detected by search A but not by B; n_B is the number of events detected by search B but not by A; and n_{AB} is the number of events detected by both the searches. The same definitions are applyed for the efficiencies $(\epsilon_A, \epsilon_B, \epsilon_{AB})$. This can be easily extendend to a general number$p$ of searches, where we have $2^p - 1$ distint combinations by which an event may be detected.

Introducing the vector notations $\vec{N}$, $\vec{\epsilon}$ and $\vec{b}$ whose components $i = \{1, ..., q\}$ refer to different searches, we have:

$$P(\vec{N}|\lambda\vec{\epsilon} + \vec{b}) = \prod_{i=1}^{p} P(N_i|\lambda\epsilon_i + b_i) \tag{5.9}$$

To set an upper limit, one must decide a priori how to rank all possible observations, so as to determine whether a given observation contains more or fewer events than some other observations. Chosen the ranking function $\zeta(n)$, the rate upper limit R_α at confidence level α is given by:

$$1 - \alpha = \sum_{\zeta(\vec{N}) \leq \zeta(\vec{n})} P(\vec{N}|\vec{\epsilon}, R_\alpha\vec{T}) \tag{5.10}$$

where $P(\vec{N}|\epsilon, R_\alpha\vec{T})$ is the prior probability of observing $\vec{N}$ given the true GW rate R_α, the livetimes of different data sets $\vec{T}$ and the detection efficiencies ϵ. The sum is taken over all N for which $\zeta(\vec{N}) \leq \zeta(\vec{n})$, i.e. over all possible outcomes $\vec{N}$ that results in fewer events than were actually observed.

For our case, we combine upper limits from the same pipelines applied to different networks which operates in disjoint time periods. So, the combination is simplified, because it is not possible to detect the same event from different networks (AB case mentioned above is not possible). So the formula assumes the simple form:

$$\mu_{90\%}^{TOT} = \frac{2.30}{\sum_i T_i\epsilon_i} \tag{5.11}$$

where i refers to the networks.

An example on the combinations of multiple pipelines for the SG3067Q9 is provided in Fig. 5.24.

The final upper limits for the SGQ9 MDC set calculated combining all the networks involved are shown in Fig. 5.25

In this work we have summarized results of the all-sky search for gravitational burst signal performed on the LIGO S5 second year and Virgo VSR1 data collected in 2006-2007, on the frequency band [1280, 6000] Hz, applying the Waveburst algorithm. The quality of the data and the sensitivity of the search algorithms have

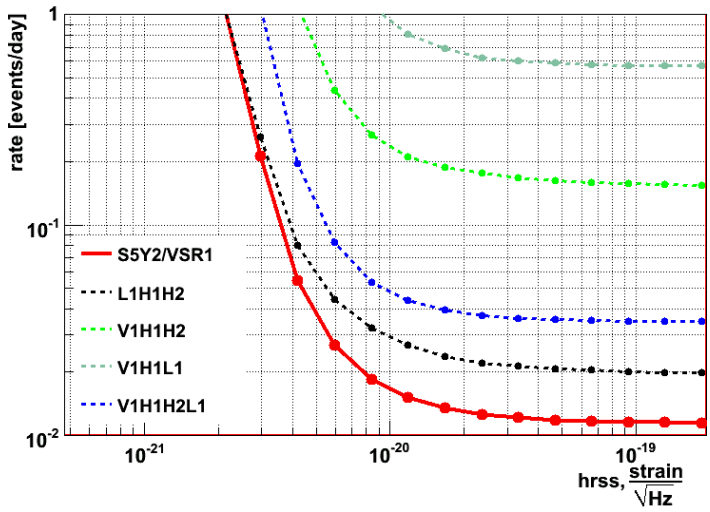

Figure 5.24: *Upper limit curves for SG3067Q9 combined from single networks using a 90% coverage.*

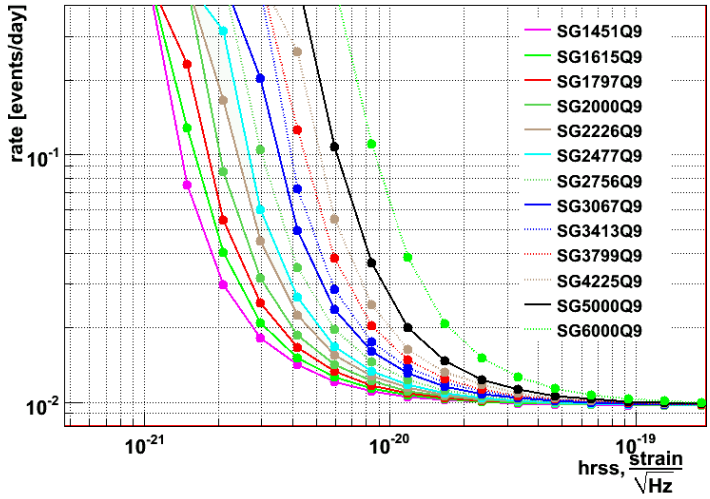

Figure 5.25: *Upper limit curves for SGQ9 combined from single networks using a 90% coverage.*

improved since similar searches of the previous LIGO S4 run [115] and even since the first year of S5 [116, 117]. This is reflected in an improved strain sensitivity with Hrss50% as low as $5.6 \cdot 10^{-22}\ 1/\sqrt{Hz}$ for certain LF waveforms, compared to

best values of $1.3{\cdot}10^{-21}\ 1/\sqrt{Hz}$ and $6.0{\cdot}10^{-22}\ 1/\sqrt{Hz}$ for S4 and S5Y1 respectively. Moreover, the S5Y2/VSR1 search also covers an extendend frequency band up to 6 kHz, and the Hrss50% for S5Y2/VSR1 is improved of about 30% respect to S5Y1 high frequency.

No plausible gravitational wave candidates have been identified in this search. Combined with the S5Y2 results, which had comparable observation time, this yields an improved upper limit on the rate of bursts of 2.2 events per year at 90% confidence for the [2000, 6000] Hz band, which is about an half with respect to the S5Y1 upper limit of 5.4 per year for the frequency band [1000, 6000] Hz.

Conclusions and future prospectives

Gravitational Wave physics is a continuos improving science field that could give many informations on the astrophysical objects and their caracteristichs that are not available from electromagnetic radiation. However, the weak interaction of the GW with the matter is an huge obstacle on the GW detection.

This is the main reason which forces the scientific community to develop high sentivity intruments and search algorithms to allows the possibility to identify with a sure reliability a GW detection.

In spite of the great effort of scientists addressed to this goal, a GW detection has not yet declared. This creates also a lack of comparison term, which makes more difficult to establish a sure method to declare a detection of a GW. Moreover, the complicated resolution of Einstein's equation applied to astrophysical processes, often requires the use of approximated numerical simulations, which cannot describe all the possible scenarios of GW generating processes. This explains the importance of introducing an algorithm which does not require particular assumptions on the expected signal we want to detect, i.e. what we have called in this thesis un-modelled search.

Moreover, these issues are strictly related to the importance of improving detector perfomances. In particular, the GW community planned to improve by an order of magnitude the sensitivity of LIGO and Virgo detectors with respect to the present ones (Advanced LIGO [19] and Advanced Virgo [118]). This would allow to increase by the same order the observable horizon and by three orders of magnitude the observable sky volume, with an expected rate of 10s/100s events per year.

In this work we have presented the **Waveburst** algorithm, focusing on the **Coherent** version, we have realised, in fact, that coherent pipelines are by far more efficient on discarding false alarms respect to the incoherent ones. Waveburst uses the wavelet properties and the likelihood ratio approach to identify significant triggers in the data stream, and it is able to adapt the search to the spectral and directional sensitivity of the network (use of *regulator*) or make particular assumptions on the waveform characteristics (use of *constrains*).

We have applied the Coherent Waveburst algorithm to the LIGO-Virgo collab-

oration detector on the first long term observation of LSC and Virgo detectors, **S5Y2/VSR1**, to search for GW transient signals in the High Frequency bandwidth ($2 \div 6$ kHz). With respect to previous analysis in the same bandwidth, this is the first coherent analysis, which is an important feature when detectors with so different directional sensitivities are in joint observation. The analysis has been performed on data from different network configurations, i.e. combinations of the four involved detectors (L1, H1, H2, V1). Each configuration is analysed in its exclusive livetime, i.e. the observation time covered by only that specific configuration of detectors in science mode. Only networks with enough live time has been analysed (L1H1H2V1, L1H1H2, V1H1H2, L1H1V1, H1H2). For each network, *background* and *simulations* tests have been performed, in such a way to give a reliable confidence of a possible detection and to give an interpretation of the results in terms of some classes of signal models. Sine Gaussian waveforms have been used as reference efficiency waveforms for the tuning of the post production parameters, where as we have tested also the efficiency detection for more interesting astrophysical signals (such as Ring Down waveforms or Supernova Collapse). No events candidate have been found. The direct scientific output is the best upper limit for this bandwidth.

The pipeline version used for the S5/VSR1 analysis had the most advanced performaces available for what concerns the capability of discriminating signals from accidental triggers. However, its performances for what concerns the reconstruction of the characteristics of an identified GW candidate were still far from optimal. We then started a research and development study to improve the capabilities of the algorithm to estimate the GW direction and we took part to the project Position Reconstruction Challenge, whose goals was to stimulate progresses on position reconstruction with LSC and Virgo detectors and compare the results from different algorithms. We achieved progresses of one order of magnitude better with respect to previous angular reconstruction. Today, error angles for a wide class of signal waveforms are typically of the order of one degree at high signal to noise ratio. However, further progresses are needed to easy the start of a multimessenger astronomy, in particular to allow a follow up of GW candidates by other astronomical instruments from X and gamma-ray satellites to wide fields optical based instruments.

The current work on reconstruction capabilities is also targeting procedures to set sky Error Region on source position coverage (probability of detecting signal). In addition, to the position reconstruction problem, we recently started to address the issue of reconstructing the waveform information from detected signals, which will be obviously as important as source location for the GW astronomy.

We are currently analyzing the second long term observation of the LSC-Virgo detectors (started on July 2009, planned to the end of 2010). The first results are being internally reviewed by the collaboration. For this transient signal search, the main structure of the algorithm has been improved in other two aspects which are relevant for the identification of signal candidates. Any position in the earth frame can now be used as reference for the delays in arrival time of the GW wavefront, so that the Likelihood Sky Map can be computed from any such reference, e.g. any detector. The time slide procedure to sample the background has been generalized so that any detector can be time shifted with respect to the others. These symmetrizations of the scheme of the pipeline allow a much better estimation of the accidental background of the searches both from what concerns statistical and systematic errors.

Appendix A

Pipeline line by line

This is a summary of the principal actions of the pipeline, following line by line the C++ macro. Some commands are applied only for simulation tests injecting MDC. In this case we use the label "(SIM)".

Inizialization
Initialization of the useful elements: mother wavelet, arrays (time, MDC, noise, detectors).
Inizialization of the network and all the elements connected to it.
Inizialization of trees[1]: mdc, livetime, bursts, variability, noise.
Reading of the file containing information about time and data for the selected job. The pipeline calculated the start and stop of the segment.
Correction of lags number to avoid time shifts greater than length segment.
NET.setTimeShifts: set parameters for time shift analysis.
(SIM) MDC log file is read.
Data frames are uploaded, the interesting bandwidth is selected using wavelet down conversion and then data are saved at a chosen wavelet decomposition level. 6 kHz Down conversion is applied at this step.
(SIM) Same procedure as the previous line for MDC data frames.
Assignations output and temporary ROOT file names.
Starting of loop for amplitude injection factors. The following commands are related to a single amplitude factor.

Data Conditioning
Uploading of Down Converted data saved previously.
(SIM) Uploading of MDC Down Converted data saved previously. Data amplitudes are multiplied for the amplitude factor. MDC is injected in the noise data stream.

[1] A tree is a ROOT object which records for each event all its characteristics (called leaves) for a simple visualization.

Line Predictor Filter and Whitening of the data stream.
A band pass filter is applied, to select the frequency band of interest previously declared (parameter file).
NET.SetVeto: set time periods which are not processed because excluded by veto file.

Inizialization of the output files

Final ROOT files are created, and finals trees are associated to them.

Start of the coeherent search

Starting of loop for TF levels. The follosing commands are related to a single decomposition level.
Uploading of time delay filters.
NET.threshold: calculates Waveburst threshold as a function of resolution and maximum delay.
NET.coherence: selects time-frequency samples.
NET.cluster: produces time-frequency cluster as a swingle time-frequency resolution.
NET.likelihood(E)
Ending of loop for TF levels.

Supercluster analysis

pwc=NET.getwc, pwc->supercluster: constructs the superclusters.
NET.events: returns number of events.

Likelihood

NET.likelihood()
NET.netcut(netRHO,r,0,1): selection of clusters based on threshold on rho.
NET.netcut(netCC,c,0,1): selection of clusters based on threshold on network correlation coefficient.
NET.events: return number of events.
(SIM) NET.printwc: save the properties of the events.

Save data in root file

Events are saved in output files.
Temporary files used for the analysis are deleted.

A.1 NET.Coherence

This functions selects the pixels with compose the clusters, tagging them with a boolean array of values. Only the selected pixels are used to compute likelihood (NET.likelihood function).

Without considering line by line this function, the procedure of selecting pixel is based on the coherence between corresponding pixels belonging to different de-

tectors. The first detector is used as reference one, and its pixels are tagged by the boolean array. The function takes one pixel of the first detector and it calculates the corresponding pixels of the other detectors associated to a certain position in the sky, using the delay time filters (Cap. 3.2.5). For a pixel of the first detector located at the time t_0, the corresponding pixels of another one refer to a time period $[t_0 - \delta t_{10},\ t_0 + \delta t_{10}]$, where δt_{10} is the maximum delay time between the two detectors. This means that the generic pixel t_0 of the reference detector is combined with:
* N_1 pixels inside $[t_0 - \delta t_{10},\ t_0 + \delta t_{10}]$ of detector 1,
* N_2 pixels inside $[t_0 - \delta t_{20},\ t_0 + \delta t_{20}]$ of detector 2,
and so on for all the detectors involved.

Each set of the type: $\{t_0,\ t_{1k},\ t_{2k},\ ...\}$ indicates a position (k) in the sky. Varying each t_{ik}, $i > 0$ inside the periods $[t_0 - \delta t_{1i},\ t_0 + \delta t_{1i}]$ we can select all the sky positions. The pixel t_0 is selected if there is at least position in the sky where the sum of square amplitude of the pixels $t_0,\ t_1,\ t_2,\ ...$ is greater than a definite threshold.

This procedure is repeated for all the pixels of all the decomposition levels.

This function is computationally heavy, so the pipeline uses some tricks to decrease the computational loading. Consider the pixel t_0, we have to calculate all the quantities $S_k = \sum_{i=0}^{N} A(t_{ik})$, where N is the number of detectors and $A(p)$ is the amplitude of the pixel p. If the $\max_k S_k >$ threshold the pixel is selected. We can write $S_k = A(t_0) + \sum_{i=1}^{N} A(t_{ik})$ so the previous condition can be written $\max_k S'_k >$ threshold $- A(t_0)$ where $S'_k = \sum_{i=1}^{N} A(t_{ik})$. This allows us to reduce computational load when applying time shift. In fact, if we have M lags of Δ time shift, the $\max_k S'_k$ has the same value for the different $t_0 + m\Delta$ $(0 \leq m \leq M)$ and we can calculate one time for all the shifts applyed.

Computational time is reduced also considering that $\max_k S'_k < \sum_{i=1}^{N} \max_k A(t_{ik})$, so we can make a pre-selection of pixels considering the disequality $\sum_{i=1}^{N} \max_k A(t_{ik}) >$ threshold. Pixels that do not survive to this disequality cannot surely be tagged as usable. This pre-selection reduces significantly the number of pixels to be selected (a factor of the order of 10^3).

A.2 NET.likelihood

This function calculates the likelihood ratio on the cluster composed by the selected pixels. The NET.likelihood function is applied twice in the pipeline, the first time (when it accepts the parameter 'E') it operates a selection on the cluster eliminating the ones which have a large discrepancy between the detector contributes.

The second times the function is applied the algorithm calculates the likelihood ratio for each event (an event is represented by the correspondind super cluster) over all the sky locations. The maximum value of the likelihood location over the sky is considered as the starting point to calculate all the parameters that we associate to the reconstruct events. The likelihood ratio is calculated according to the chosen regulator and constrain (See Chapter 3.1.4 and Appendix B).

Appendix B

Relation between a general frame and the DPF

We have introduced the DPF in the section 3.1.2, demonstrating that in this coordinate the Likelihood functional assumes a simple form. However, it is not necessary to transpose all the vectors in this frame, but it is possible to calculate the Likelihood functional and all the interesting quantities also in the general frame. We remember now useful consideration.

- In the general frame, the vectors f_+ and $f_\times$ are not orthogonal.
- The transformation between the general frame and DPF are given by:

$$\begin{aligned} f'_{k+} &= f_{k+}Cos(\gamma) + f_{k\times}Sin(\gamma) \\ f'_{k\times} &= -f_{k+}Sin(\gamma) + f_{k\times}Cos(\gamma) \end{aligned} \tag{B.1}$$

  where

$$\begin{aligned} Cos(2\gamma)|A_\sigma|^2 &= (f_+^2 - f_\times^2) \\ Sin(2\gamma)|A_\sigma|^2 &= 2(f_+ \cdot f_\times) \end{aligned} \tag{B.2}$$

- We are more interested in direction of vectors than in module, because we can always normalize them. This means that sometimes we will calculate the quantity αv instead of v if it is easier computable.
- The Maximum Likelihood is given by:

$$L = \frac{(X \cdot u)^2}{|u|^2} \tag{B.3}$$

  where u is defined in the $(f'_+,\, f'_\times)$ plane by regulators.
- $(f'_+,\, f'_\times)$ and $(f_+,\, f_\times)$ are the same plane, because the f'_+ and $f'_\times$ vectors are linear combinations of f_+ and $f_\times$.

From these considerations we can obtain the formula of the Likelihood functional in the general frame, calculating the expression of vector u in the general frame for the different regulators considered.

The vector u can be expressed as a linear combination of the two vectors f_+ and $f_\times$:

$$u = u_c f_+ + u_s f_\times \tag{B.4}$$

so in the following considerations we will find the expression of u_c and u_s for the different regulators definitions.

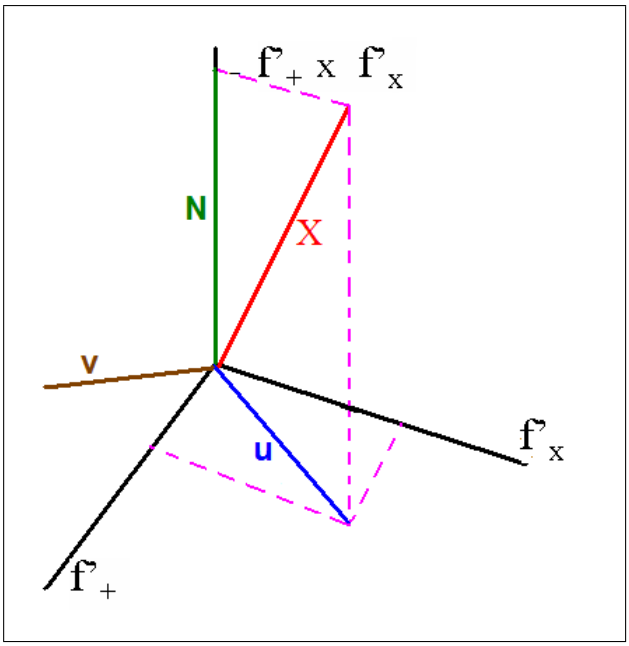

Figure B.1: *Dominant Polarization Frame, v is a vector perpendicolar to u in the plane* $(f'_+, f'_\times)$

B.1 Weak regulator

The **weak** regulator consider u as the projection of X in the $(f_+, f_\times)$ plane. Because f_+ and $f_\times$ are not orthogonal, we cannot write $u = \frac{u \cdot f_+}{|f_+|} e_+ + \frac{u \cdot f_\times}{|f_\times|} e_\times$[1] as we can do in the DPF. We can avoid this problem using the vector product.

[1] e_+ and $e_\times$ are versor with the same direction as f_+ and $f_\times$.

Referring to Fig. B.1, we consider the vector $f_+ \times f_\times$, which describe the sub-space orthogonal to the plane (f_+, $f_\times$). Then we define the vector v

$$v = X \times (f_+ \times f_\times) \tag{B.5}$$

This vector is on the plane (f_+, $f_\times$) and it is orthogonal to X. So, if we define the vector

$$u = -v \times (f_+ \times f_\times) \tag{B.6}$$

u is on the plane (f_+, $f_\times$) (because it is orthogonal to $f_+ \times f_\times$) and (apart from the module) it is the projection of X (because it is othogonal to v, which is orthogonal to X). This is the vector we are interested in, except for the module, but we can disregard about it.

Mathematically, we use the property of triple vector product:

$$a \times (b \times c) = b(a \cdot c) - c(a \cdot b) \tag{B.7}$$

we have:

$$u = -v \times (f_+ \times f_\times) = -f_+(v \cdot f_\times) + f_\times(v \cdot f_+) \tag{B.8}$$

Comparing this equality with $u = f_+ u_c + f_\times u_s$ we have: $u_c = -v \cdot f_\times$ and $u_s = v \cdot f_+$. Remebering the definition of v:

$$v = X \times (f_+ \times f_\times) = f_+(X \cdot f_\times) - f_\times(X \cdot f_+) = f_+ X_\times - f_\times X_+ \tag{B.9}$$

where we have defined the quantities: $X_+ = X \cdot f_+$ and $X_\times = X \cdot f_\times$. Then:

$$\begin{aligned} -v \cdot f_\times &= -(f_+ X_\times - f_\times X_+) \cdot f_\times = X_+(f_\times \cdot f_\times) - X_\times(f_+ \cdot f_\times) \\ v \cdot f_+ &= (f_+ X_\times - f_\times X_+) \cdot f_+ = X_\times(f_+ \cdot f_+) - X_+(f_\times \cdot f_+) \end{aligned} \tag{B.10}$$

Finally:

$$\begin{aligned} u_c(\text{weak}) &= X_+(f_\times \cdot f_\times) - X_\times(f_+ \cdot f_\times) \\ u_s(\text{weak}) &= X_\times(f_+ \cdot f_+) - X_+(f_\times \cdot f_+) \end{aligned} \tag{B.11}$$

B.2 Soft regulator

The **soft** regulator defines $u = (X \cdot f'_+)f'_+ + (X \cdot f'_\times)f'_\times$. So we can write:

$$\begin{aligned}
u &= [X \cdot (f_+ Cos(\psi) + f_\times Sin(\psi))](f_+ Cos(\psi) + f_\times Sin(\psi)) + \\
&+ [X \cdot (-f_+ Sin(\psi) + f_\times Cos(\psi))](-f_+ Sin(\psi) + f_\times Cos(\psi)) = \\
&= [X_+ Cos(\psi) + X_\times Sin(\psi)](f_+ Cos(\psi) + f_\times Sin(\psi)) + \\
&+ [-X_+ Sin(\psi) + X_\times Cos(\psi)](-f_+ Sin(\psi) + f_\times Cos(\psi)) = \\
&= [X_+ Cos^2(\psi) + X_\times Cos(\psi)Sin(\psi) + X_+ Sin^2(\psi) - X_\times Cos(\psi)Sin(\psi)]f_+ + \\
&+ [X_+ Cos(\psi)Sin(\psi) + X_\times Sin^2(\psi) - X_+ Cos(\psi)Sin(\psi) + X_\times Cos^2(\psi)]f_\times = \\
&= X_+ f_+ + X_\times f_\times
\end{aligned} \tag{B.12}$$

So, the definition of u for the soft regulator is:

$$\begin{aligned}
u_c(\text{soft}) &= X_+ \\
u_s(\text{soft}) &= X_\times
\end{aligned} \tag{B.13}$$

B.3 Hard regulator

The **hard** regulator assumes that u has the same direction as f'_+ in the DPF. So, $u = f'_+$, and from Eq. 3.17 we have that $f'_+ = f_+ Cos(\gamma) + f_\times Sin(\gamma)$. Because we are interested only in direction, we calculate the quantity $Cos(\gamma)u$ instead of u:

$$Cos(\gamma)u = f_+ Cos^2(\gamma) + f_\times Sin(\gamma)Cos(\gamma) = f_+ \frac{1}{2}[1 + Cos(2\gamma)] + f_\times \frac{1}{2} Sin(2\gamma) \tag{B.14}$$

and from Eq. 3.21 we have:

$$Cos(\gamma)u = f_+ \frac{1}{2}\left[1 + \frac{f_+^2 - f_\times^2}{|A_\sigma^2|}\right] + f_\times \frac{f_+ \cdot f_\times}{|A_\sigma^2|} \tag{B.15}$$

So, multipling by $|A_\sigma^2|$ and discarding about the factors on the left part we have the expression for u_c and u_s:

$$\begin{aligned}
u_c(\text{hard}) &= \tfrac{1}{2}|A_\sigma^2| + \tfrac{1}{2}(f_+^2 - f_\times^2) \\
u_s(\text{hard}) &= f_+ \cdot f_\times
\end{aligned} \tag{B.16}$$

B.4 Mild regulator

Finally the **mild** regulator defines $u = f'_+ \pm f'_\times$. We can repeat the same procedure for the hard regulator.

$$\begin{aligned} Cos(\gamma)u &= Cos(\gamma)(f'_+ \pm f'_\times) = \\ &= (f_+ Cos^2(\gamma) + f_\times Cos(\gamma)Sin(\gamma)) \pm (-f_+ Cos(\gamma)Sin(\gamma) + f_\times Cos^2(\gamma)) = \\ &= f_+(Cos^2(\gamma) \mp Cos(\gamma)Sin(\gamma)) + f_\times(Sin(\gamma)Cos(\gamma) \pm Cos^2(\gamma)) = \\ &= f_+ \left[\frac{1}{2}\left(1 + \frac{f_+^2 - f_\times^2}{|A_\sigma^2|}\right) \mp \frac{f_+ \cdot f_\times}{|A_\sigma^2|}\right] + f_\times \left[\frac{f_+ \cdot f_\times}{|A_\sigma^2|} \pm \frac{1}{2}\left(1 + \frac{f_+^2 - f_\times^2}{|A_\sigma|^2}\right)\right] \end{aligned} \tag{B.17}$$

Multipling for $|A_\sigma^2|$ and discarding for the common factor, we obtain the final form:

$$\begin{aligned} u_c(\text{mild}) &= \tfrac{1}{2}(|A_\sigma^2| + (f_+^2 - f_\times^2)) \mp (f_+ \cdot f_\times) \\ u_s(\text{mild}) &= (f_+ \cdot f_\times) \pm \tfrac{1}{2}(|A_\sigma^2| - (f_+^2 - f_\times^2)) \end{aligned} \tag{B.18}$$

The $\pm$ sign depends on what quadrants of the f_+,$f_\times$ plane the X vector is projected. So, if the projection of X is on the first or third quadrant the sign is $+$, otherwise this sign is $-$. This conditions can be expressed using the inequalities on the Tab. B.1:

Condition	Sign
$(X \cdot f'_+)(X \cdot f'_\times) > 0$	+
$(X \cdot f'_+)(X \cdot f'_\times) < 0$	-

Table B.1: *Conditions for the choice of $\pm$ sign of the mild regulator*

B.5 Intermediate regulators

It is possible to choose intermediate situations between regulators using variable parameters. For instance in the definition of mild regulators: $u_c(\text{mild}) = 1/2|A_\sigma^2| + 1/2(f_+^2 - f_\times^2) \mp (f_+ \cdot f_\times)$ we can define a parameter called *mild* in the following way: $u_c(\text{mild}) = 1/2|A_\sigma^2| + 1/2(f_+^2 - f_\times^2) \mp \text{mild}(f_+ \cdot f_\times)$. Changing the value of the parameter *mild* from 0 to 1 we can use an intermediate regulator from mild (mild=1) to hard (mild=0) definitions. Generally we can define an single definition

of u_c and u_s including all the regulators:

$$\begin{aligned}
u_c(\text{all}) &= \text{soft}X_+ + \\
&+ \text{mild}(f_+ \cdot f_\times)\text{gg} + \\
&+ \frac{\text{hard}}{(f_+^2 + f_\times^2)/2 + |A_\sigma^2|}\left[X_+\left(|A_\sigma^2| + \frac{f_+^2 - f_\times^2}{2}\right) + X_\times(f_+ \cdot f_\times)\right] \\
u_s(\text{all}) &= \text{soft}X_\times + \\
&+ \text{mild}\left(|A_\sigma^2| + \frac{f_+^2 - f_\times^2}{2}\right)\text{gg} + \\
&+ \frac{\text{hard}}{(f_+^2 + f_\times^2)/2 + |A_\sigma^2|}\left[X_\times\left(|A_\sigma^2| - \frac{f_+^2 - f_\times^2}{2}\right) + X_+(f_+ \cdot f_\times)\right]
\end{aligned} \tag{B.19}$$

where gg is defined as:

$$\begin{aligned}
\text{gg} = &\frac{1}{\left(|A_\sigma|^2 + \frac{f_+^2 - f_\times^2}{2}\right)^2 + (f_+ \cdot f_\times)^2} \cdot \\
&\left\{\left[X_\times\left(|A_\sigma^2| + \frac{f_+^2 - f_\times^2}{2}\right) - X_+(f_+ \cdot f_\times)\right] \pm \left[X_+\left(|A_\sigma^2| + \frac{f_+^2 - f_\times^2}{2}\right) + X_\times(f_+ \cdot f_\times)\right]\right\}
\end{aligned} \tag{B.20}$$

Using intermediate values of the parameters we can use superimposition of regulators, in particular we can define a new parameter (δ) and varying this we can choose a regulator o a superimposition between regulators (Tab B.2).

δ **range**	**Regulator**	**soft**	**mild**	**hard**
δ = -1	soft	1	0	-1-δ=0
-1 < δ < 0	soft-weak	1	0	-1-δ
δ = 0	weak	1	δ=0	δ-1=-1
0 < δ < 1	weak-mild	1	δ	δ-1
δ = 1	mild	1-hard=1	1-hard=1	δ-1=0
1 < δ < 2	mild-hard	1-hard	1-hard	δ-1
$\delta \geq 2$	hard	0	0	1

Table B.2: *Relationship between δ and parameter values to obtain regulators from Eq. B.19. Red cases refer to pure regulator definition, blue cases refer to superimpositions between regulators.*

Appendix C

Energy disbalance approximations

Equation 3.49 is an approximation of 3.47 for $\psi \approx 0$. Reminding the quantities involved:

$$\begin{array}{c} u_k = u_{k0}Cos(\psi) + v_{k0}Sin(\psi) \\ v_k = \frac{\partial u}{\partial \psi} = -u_{k0}Sin(\psi) + v_{k0}Cos(\psi) \\ q_k = X_k u_k - u_k^2 h \\ p_k = -\frac{\partial q_k}{\partial \psi} = (2u_k h - X_k)v_k + u_k^2(X \cdot v)/|u|^2 \end{array} \tag{C.1}$$

where the index k refers to a certain detector, and the scalar product $(\ \cdot\ )$ substitutes the sum between indexes k, as we have previously introduced.

Now we approximate the definitions of p_k and q_k for small angles ψ. We remind that in the limit $\psi \to 0$, then $Sin^n(\psi) \sim 0$ for $n \geq 2$:

$$\begin{array}{c} Sin^2(\psi) \sim 0 \\ Cos^2(\psi) = 1 - Sin^2(\psi) \sim 1 \end{array} \tag{C.2}$$

Obviously the module of u remains constant changing the value of ψ angle, in fact this is only a rotation on the plane. Matemathically:

$$\frac{\partial |u|^2}{\partial \psi} = \frac{\partial (u \cdot u)}{\partial \psi} = 2(u \cdot v) = 0 \tag{C.3}$$

Considering the definition of q_k and using the previous approximations:

$$
\begin{aligned}
q_k &= X_k u_k - u_k^2 h = \\
&= X_k(u_{k0}Cos(\psi) + v_{k0}Sin(\psi)) + \\
&- (u_{k0}Cos(\psi) + v_{k0}Sin(\psi))^2 \frac{X \cdot (u_0 Cos(\psi) + v_0 Sin(\psi))}{|u|^2} \approx \\
&= X_k u_{k0} Cos(\psi) + X_k v_{k0} Sin(\psi) + \\
&- (u_{k0}^2 Cos^2(\psi) + 2u_{k0}v_{k0}Cos(\psi)Sin(\psi)) \frac{(X \cdot u_0)Cos(\psi) + (X \cdot v_0)Sin(\psi)}{|u_0|^2} = \\
&= Cos(\psi)\left[X_k u_{k0} - u_{k0}^2 \frac{X \cdot u_0}{|u_0|^2}\right] + \\
&+ Sin(\psi)\left[X_k v_{k0} - u_{k0}^2 \frac{X \cdot v_0}{|u_0|^2} - 2u_{k0}v_{k0}\frac{X \cdot u_0}{|u_0|^2}\right] = \\
&= q_{k0}Cos(\psi) - p_{k0}Sin(\psi)
\end{aligned}
\tag{C.4}
$$

The approssimation for p_k is calculated from its definition:

$$
\begin{aligned}
p_k = -\frac{\partial q_k}{\partial \psi} &\sim -\frac{\partial}{\partial \psi}(q_{k0}Cos(\psi) - p_{k0}Sin(\psi)) \\
&= q_0 Sin(\psi) + p_0 Cos(\psi)
\end{aligned}
\tag{C.5}
$$

Bibliography

[1] F. Acernese et al., *The Virgo status*, Class. Quantum Grav. 23 (2006) S635

[2] B.P. Abbott et al., *LIGO, the Laser Interfometry Gavirational-Wave Observatory*, Rep. Prog. Phys.72 (2009) 076901

[3] H. Lück et al., *Status of the GEO600 detector*, Class. Quant. Grav. 23 S71

[4] R. Takahashi, *Status of TAMA 300*, Class. Quant. Grav 21 (2004) S403

[5] K. S. Thorne, *Gravitational waves from compact bodies*, in *Proc. of IAU Symposiusm 165, Compact Star in Binaries*, (1995)

[6] K. S. Thorne, in *300 Years of Gravitations*, Hawking and W. Israel eds, Cambridge Univ. Press (1987)

[7] J. Hartle, *Gravity, an introduction to Einstein's general relativity*, Addison Wesley, 2003.

[8] C. D. Ott, *The gravitational-wave signature of core-collapse supernovae*, Class Quantum Grav. 26, 063001 (2009)

[9] E. Baron and J. Copperstein, Astrophys. J. 353, 597 (1990)

[10] H. A. Bethe, *Supernova mechanism*, Rev. Mod. Phys. 62, 801 (1990)

[11] H-T Janka et al., *Theory of core-collapse supernovae*, Phys Rep 442, 38 (2007)

[12] J.M LeBlanc and J.R. Wilson, Astrophys J. 161, 541 (1970)

[13] H Swai et al., Astrophys J 672, 465 (2008)

[14] A. Burrows et al., *Multi-dimensional explorations in supernova theory*, Phys. Rep. 442, 23 (2007)

[15] S. van den Bergh and G.A. Tammanm, *Galactic and Extragalactic Supernova Rates*, Ann. Rev. Astron. Astroph. 29, 363 (1991)

[16] S. Andoet al, Astron. Astroph. 246, 47 (1991)

[17] L. Baiotti and L. Rezzolla, *Challenging the Paradigm of Singularity Excision in Gravitational Collapse*, Phys. Rev. Lett. 97, 141101 (2006)

[18] L. Baiotti et al., *On the gravitational radiation from the collapse of a neutron stars to rotating black holes*, Class. Quantum Grav. 24, S187 (2007)

[19] P. Fritschel, *Second generation instruments for the Laser Interferometer Gravitational Wave Observatory (LIGO)*, Proceedings of SPIE, Vol. 4856, p. 282 (2003)

[20] C. D. Ott, *Stellar iron core collapse in 3+1 general relativity and the gravitational wave signature of core-collapse supernovae*, PhD Thesis Universität Potsdam, Germany

[21] R. Oechslin and H-J Janka, *Gravitational Waves from Relativistic Neutron-Star Mergers with Microphysical Equations of State*, Phys. Rev. Lett. 99, 121102 (2007)

[22] K. Kiuchi et al., *Longterm general relativistic simulation of binary neutron stars collapsing to a black hole*, submitted to Phys. Rev. D, arXiv:0904.4551

[23] A. Buonanno et al., *Detection template families for gravitational waves from the final stages of binary-black-hole inspirals: Nonspinning case*, arXiv:gr-gc/0205122v3 (2006)

[24] N. Andersson e K.D. Kokkotas, *Gravitational-wave astronomy: the high-frequency window*, General Relativity and Quantum Cosmology, 2004.

[25] V. Ferrari, G. Minuitti and J.A. Pons, *Gravitational waves from newly born, hot neutron stars*, Mon. Not. R. Astron. Soc. 2006.

[26] V. Kalogera et al. Astrophys J. 601, L179 (2004)

[27] C. Kim et al., *The Probability Distribution of Binary Pulsar Coalescence Rates.*, Astrophys J. 584, 985 (2003)

[28] K. T. Inoue and T. Tanaka, *Gravitational Waves from Sub-Lunar-Mass Primordial Black-Hole Binaries: A New Probe of Extradimensions*, Phys. Rev. Lett. 91, 021101 (2003)

[29] S. E. Woosley, Astrophys. J. 405, 273 (1993)

[30] B. Paczynski, Astrophys. J. 494, L45 (1998)

[31] A. I. MacFadyen and A. I. Woosley, *Collapsars: Gamma-Ray Bursts and Explosion in "Failed Supernovae"*, Astrophys. J. 524, 262 (1999)

[32] S. I. Blinnikov et alt., Sov. Astron. Lett. 10, 177 (1984)

[33] J. Goodman, Astrophys. J. 308, L47 (1986)

[34] J. Goodman et al., Astrophys. J. 314, L7 (1987)

[35] D. Eichler, Mon. Not. Roy. Astron. Soc. 335, 883 (2002)

[36] R. F. Stark and T. Piran, *Gravitational-Wave Emission from Rotating Gravitational Collapse*, Phys. Rev. Lett. 55, 891 (1985)

[37] B. F. Schultz, Class. Quantum Grav. 16, A131 (1999)

[38] J. G. Jernigan, *KiloHerts QPO and Gravitational Wave Emission as the Signature of the Rotation and Precession of a LMXB Neutron Star Near Breakup*, AIP Conf. Proc. 586, 805 (2001)

[39] H. J. Mosquera Cuesta and D. M. Gonzales, *Bursts of Gravitational Radiation from Superconducting Cosmic Strings and the Nuetrino Mass Spectrum*, Phys. Lett. B 500, 215-221 (2001)

[40] www.auriga.lnl.infn.it

[41] Ott C, Burrows A, Dessart L and Livne E, *Physical Review Letters*, **96** 201102, 2006.

[42] www.virgo.infn.it

[43] www.ligo.caltech.edu

[44] www.geo600.uni-hannover.de

[45] http://tamago.mtk.nao.ac.jp

[46] http://www.gravity.uwa.edu.au

[47] http:///www.lisa.jpl.nasa.gov

[48] R. V. Pound, *Electronic frequency stabilization of microwave oscillators*, Rev. Sci. Instrum. 17, 490-505 (1946)

[49] R. W. P. Drever et al., *Laser phase and frequency stabilization using an optical resonator*, Appl. Phys. B 31, 97-105 (1983)

[50] Virgo Collaboration, *Extending the virgo gravitational wave detection band down to a few Hz: metal blade springs and magnetic antisprings*, Nucl. Instrum. and Meth. in Phys. Red. A, 394:397-408 (1997)

[51] J. N. Brune and J. Oliver, *The seismic noise of the earth's surface*, Bull. Seism. Soc. Am. 49 (1959)

[52] Virgo collaboration, *The virgo interferometer*, Class. Quantum Grav. 14, 1461 (1997)

[53] H. B. Callen and T. A. Welton, Phys. Rev. 83 (1951)

[54] C. Cagnoli, L. Gammaiotni, J. Kovalik, F. Marchesoni and M. Punturo, *The Virgo sensitivity curve*, virgo note, Perugia March 1997

[55] C. Corda, *Generalized gauge-invariance for gravitational waves*, arXiv:0706.2412v1 (2007)

[56] S.Klimenko e G. Mitselmakher, *A Wavelet method for detection of gravitational wave bursts*, Class. Quantum Grav. 21 (2004) S1819-S1930.

[57] S. Klimenko et al., *A coherent method for detection of gravitational wave bursts*, Class. Quantum Grav. 25:S114029 (2008)

[58] S.Chatterji *Multiresolution techniques for the detection of gravitational-wave bursts*, Class. Quantum Grav. 21, (2004) S1809

[59] A. Clapson et al., *EGC: a time-frequency augmented template-based method for gravitational wave burst search in ground based interferometers*, Class. Quant. Grav 25:S035002 (2008)

[60] M.Bignotto et al., *A cross-correlation method to search for gravitational wave burts with AURIGA-Virgo*, Class Quantum Grav. 25, 114046 (2008)

[61] J. Markowitz et al., *Gravitational wave burst source direction estimation using time and amplitude information*, Phys. Rev. D 78 122003 (2008)

[62] F. Cavalier et al., *Reconstruction of source location in a network of gravitational wave interferometric detectors*, Phys. Rev. D. 78:082004 (2006)

[63] S. Klimenko et al., *Waveburst, version 5*, Technical Note LIGO-T050222-00-Z, Sept 25, 2005.

[64] S. Klimenko et al., *Constraint likelihood analysis for a network of gravitational wave detectors*, Phys. Rev. D 72, 122002 (2005)

[65] http://root.cern.ch

[66] S. Klimenko, I. Yakusin and G. Mitselmakher, *Waveburst*, LIGO note T040040-00-Z

[67] S. Klimenko, I. Yakusin and G. Mitselmakher, *Waveburst, S5 version*, LIGO note T060112-00-Z

[68] S. Klimenko, I. Yakusin and G. Mitselmakher, *Waveburst, version 5*, LIGO note T050222-00-Z

[69] S. Mallat, *A wavelet tour of signal processing*, Academic Press, 1998.

[70] R. A. Mercer and S. Klimenko *Visualising gravitational-wave event candidates using the Coherent Event Display*, Class. Quantum Grav. 25, (2008) 184025

[71] R. A. Mercer and S. Klimenko, *The Coherent Event Display*, LIGO note T080077-00-Z, 2008

[72] J. Sylvestre, *Prospects for the detection of electromagenic counterparts to gravitational wave events*, Astrophys. J. 591, 1152-6 (2003)

[73] L. Li and B. Pacyznski, Astrophys. J. 507, L59-62 (1998)

[74] H. T. Janka and M. Ruffert, *Detectable Signals from Mergers of Compact Stars*, ASP Conf. Ser. 263, 333 (2002)

[75] S. Rosswog, *Mergers of Neutron Star-Black Hole Binaries with small Mass Ratios: Nucleosynthesis, Gamma-Ray Bursts, and Electromagnetic Transients*, Astrophys. J. 634, 1202-13 (2005)

[76] J. Kanner et al., *LOOC UP: locating and observing optical counterparts to gravitational wave bursts*, Class. Quantum. Grav. 25, 184034 (2008)

[77] N. Gehrels et al., *Gamma-ray bursts in the Swift era*, New Journal of Physics 9, 37 (2007)

[78] E. Nakar, *Short Hard gamma-ray Bursts*, Physics Reports 442, 166-236 (2007)

[79] http://www.ligo.caltech.edu/docs/ScienceDocs/G/G080465-02/G080465-02.pdf

[80] A. Sadowsky et al., *Simulations of Dynamo Action in Fully Convective Stars*, Astrophys. J. 676, 1162-69 (2008)

[81] J. s. Blooms et al, *Astro2010 Decadal Survey Whitepaper, Coordinated Science in the Gravitational and Electromagnetic Skies*, arXiv:0902.1527 (2009)

[82] http://gwdaw12.mit.edu

[83] http://cgwa.phys.utb.edu/gwdaw13/

[84] B. Abbott et al., *Implications for the Origin of GRB 070201 from LIGO Observations*, Astrophys. J. 681, 1419 (2008)

[85] B. Abbott et al., *Search for gravitational waves associated with 39 gamma-ray bursts using data from the second, third, and fourth LIGO runs*, Phys. Rev. D 77, 062004 (2008)

[86] B. Abbott et al., *Search for Gravitational-Wave Bursts from Soft Gamma Repeaters*, Phys. Rev. Lett. 101, 211102 (2008)

[87] https://www.lsc-group.phys.uwm.edu/twiki/bin/view/Bursts/Online

[88] https://www.lsc-group.phys.uwm.edu/twiki/bin/view/Bursts/PosRec

[89] http://www.rotse.net

[90] http://tarot.obs-hp.fr

[91] http://msowww.anu.edu.au/skymapper

[92] http://www.mpa-garching.mpg.de/~grb07/Presentations.Kulkarni.pdf

[93] http://hepwww.physics.yale.edu/quest

[94] http://grb.fuw.edu.pl/

[95] http://www.lowell.edu./users/elgb/loneos_disc.html

[96] http://www.zt.science.uwa.edu.au/

[97] http://www.nasa.gov/swift

[98] Y. Gursel and M. Tinto, *Near optimal solution to the inverse problem for gravitational-wave burst*, Phys. Rev. D 40, 3884-3938 (1989)

[99] http://tamago.mtk.nao.ac.jp

[100] A. C. Searle et al., *Robust Bayesian detection of unmodelled burst*, Class. Quantum Grav. 25, 114038 (2008)

[101] S. Klimenko et al., *Coherent Waveburst, wat version 4.7.0*, Technical Note LIGO-T080011-00-Z, May 8, 2008.

[102] J. Betswieser et al., *S5 V3 h(t) review and validation*, Technical note LIGO-T080242-00-R, Sept, 23, 2008

[103] F. Marion et al., *h(t) reconstruction for VSR1; Version 2 and 3*, VIR-078A-08, Sept 19, 2008

[104] http://www.vdb.virgo.infn.it/VDB/main.php

[105] M. A. Bizouard et al. *Science, injection, and data quality segments for Virgo*, VIR-XXX-08, Jan 14, 2008

[106] http://www.lsc-group.phys.uwm.edu/twiki/bin/view/Bursts/S5y2VSR1vetosegs

[107] http://wwwcascina.virgo.infn.it/DataAnalysis/VDBdoc/BURST/VSR1_DQ_BURST_LIST_20080726.html

[108] http://lancelot.mit.edu/∼cadonati/S5/DQ-Summary-v5/Summary.html

[109] http://emvogil-3.mit.edu/∼bhughey/high_freq_search/v3v4/v3v4strainamplitude.html

[110] http://www.gravity.phys.uwm.edu/cgi-bin/cvs/viewcvs.cgi/matapps/src/simulation/BurstMDC/documentation/?cvsroot=lscsoft

[111] https://gravity.psu.edu/∼psurg/sims/BurstMDC/Generated_Burst_MDCs.html#HSG7_R3

[112] L. Baiotti et al., *On the gravitational radiation from the collapse of neutron stars to rotating black holes*, Class. Quant. Grav 24 (2006), S187-S206

[113] G. J. Feldman and R. D. Cousins, *Unified approach to the classical statistical analysis of small signals*, Phys. Rev. D 57, 3873 (1998)

[114] P. Sutton, *Upper Limits from Counting Experiments with Multiple Pipelines*, arXiv:0905.4089 (2009)

[115] B. P. Abbott et al., *Search for gravitational-wave bursts in LIGO data from the fourth science run*, Class. Quantum Grav. 24, 5343 (2007)

[116] B. P. Abbott et al., *Search for gravitational-wave bursts in the first yearf o the fifth LIGO science run*, Phys. Rev. D 80, 102001 (2009)

[117] B. P. Abbott et al., *Search for high frequency gravitational-wave bursts in the first calendar year of LIGO's fifth science run*, Phys. Rev. D 80, 102002 (2009)

[118] G. Losurdo, *The next generation of GW detectors*, talk at the AAS conference, Austin (2008)